

Genetic Engineering

Volume 1

Principles, Mechanism, and Expression



Tariq Ahmad Bhat | Jameel M. Al-Khayri
Editors



CRC Press
Taylor & Francis Group

APPLE ACADEMIC PRESS

GENETIC ENGINEERING

Volume 1

Principles, Mechanism, and Expression

Genetic Engineering, 2-volume set

ISBN: 978-1-77491-271-3 (hbk)

ISBN: 978-1-77491-272-0 (pbk)

Genetic Engineering, Volume 1: Principles Mechanism, and Expression

ISBN: 978-1-77491-267-6 (hbk)

ISBN: 978-1-77491-268-3 (pbk)

ISBN: 978-1-00337-826-6 (ebk)

Genetic Engineering, Volume 2: Applications, Bioethics, and Biosafety

ISBN: 978-1-77491-269-0 (hbk)

ISBN: 978-1-77491-270-6 (pbk)

ISBN: 978-1-00337-827-3 (ebk)

GENETIC ENGINEERING

Volume 1

Principles, Mechanism, and Expression

Edited by

Tariq Ahmad Bhat, PhD

Jameel M. Al-Khayri, PhD



First edition published 2024

Apple Academic Press Inc.
1265 Goldenrod Circle, NE,
Palm Bay, FL 32905 USA
760 Laurentian Drive, Unit 19,
Burlington, ON L7N 0A4, CANADA

CRC Press
6000 Broken Sound Parkway NW,
Suite 300, Boca Raton, FL 33487-2742 USA
4 Park Square, Milton Park,
Abingdon, Oxon, OX14 4RN UK

© 2024 by Apple Academic Press, Inc.

Apple Academic Press exclusively co-publishes with CRC Press, an imprint of Taylor & Francis Group, LLC

Reasonable efforts have been made to publish reliable data and information, but the authors, editors, and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors, editors, and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged, please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, access www.copyright.com or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. For works that are not available on CCC please contact mpkbookspermissions@tandf.co.uk

Trademark notice: Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

Library and Archives Canada Cataloguing in Publication

.....
CIP data on file with Canada Library and Archives
.....

Library of Congress Cataloging-in-Publication Data

.....
CIP data on file with US Library of Congress
.....

ISBN: 978-1-77491-267-6 (hbk)
ISBN: 978-1-77491-268-3 (pbk)
ISBN: 978-1-00337-826-6 (ebk)

Dedication

This book is dedicated to:



Fatima bint Muhammad Al-Fihriya Al-Qurashiya

An Arab Muslim woman who is attributed with founding the oldest existing and continually operating, and first degree-awarding educational institution for natural sciences in the world, the University of al-Qarawiyyin Fez, Morocco, in 859 CE, she is also known as “Umm al-Banīn.”

(Source: wikipedia.org)



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

About the Editors



Tariq Ahmad Bhat, PhD

*Lecturer on Botany, Department of Education,
Govt. of Jammu and Kashmir, India*

Tariq Ahmad Bhat, PhD, with 18 years of teaching experience, is a lecturer on botany in the Department of Education, Govt. of Jammu and Kashmir, India, and is engaged in active research of molecular biology, cell biology, mutation breeding, and genetic improvement of legumes and medicinal plants. He has published 10 international books, 90 research papers, review articles, and book chapters. He has participated in 50 conferences, training programs, and workshops. He is one of the founder faculty members of the Chief Minister's Super 50 NEET Programme in Jammu and Kashmir, India. He is serving as the District Coordinator Anantnag of a prestigious project on medicinal plants under financial assistance of the National Medicinal Plants Board (NMPB), New Delhi (Ministry of AYUSH), India. The Government of India conferred on him the Best innovative Science Teacher Award 2014 in recognition of his meritorious services. Dr. Bhat has received his MSc and PhD from AMU, Aligarh India.



Jameel M. Al-Khayri, PhD

*Professor of Plant Biotechnology,
Department of Agricultural Biotechnology,
College of Agriculture and Food Sciences,
King Faisal University, Saudi Arabia*

Jameel M. Al-Khayri, PhD, is Professor of Plant Biotechnology at the Department of Agricultural Biotechnology, College of Agriculture and Food Sciences, King Faisal University, Saudi Arabia.

He has dedicated his research efforts on date palm biotechnology for the last three decades. He has published over 60 research articles and reviews in international journals in addition to 30 book chapters. Dr. Al-Khayri is editor of several special issues of international journals on date palm, biotechnology, and sustainable agriculture under abiotic and biotic stress. He is editor of 15 Springer reference books, including *Date Palm Biotechnology*, *Date Palm Genetic Resources and Utilization* (2 volumes), *Date Palm Biotechnology Protocols* (2 volumes), and *Advances in Plant Breeding Strategies* (9 volumes). He is a member of the editorial board and reviewers' panels of several international journals. He has participated in the organizing and scientific committees of international scientific conferences and contributed over 50 research presentations. In addition to teaching, graduate students advising, and conducting funded research projects, he has held administrative posts as Assistant Director of the Date Palm Research Center, Head of Department of Plant Biotechnology, and Vice Dean for Development and Quality Assurance. Dr. Al-Khayri is an active member of the International Society for Horticultural Science and the Society for In Vitro Biology and serves as the National Correspondent of the International Association of Plant Tissue Culture and Biotechnology. He served as a member of Majlis Ash-Shura (Saudi Arabia Legislative Council) Fifth Session. Currently, he maintains an active research program on date palm focusing on genetic transformation, secondary metabolites, and in vitro mutagenesis to enhance tolerance to abiotic and biotic stress. He is interested in the role of biotechnology in enhancing food security and the impact of global climate change on agriculture. Dr. Al-Khayri earned a BS in Biology from the University of Toledo and an MS in Agronomy and PhD in Plant Science from the University of Arkansas, USA.

Contents

<i>Contributors</i>	<i>xi</i>
<i>Abbreviations</i>	<i>xv</i>
<i>Acknowledgment</i>	<i>xix</i>
<i>Foreword</i>	<i>xxi</i>
<i>Preface</i>	<i>xxiii</i>
1. Concepts of Genetic Engineering	1
Mohammad Amin Lone and Anzar A. Shah	
2. Enzymes of Genetic Engineering	23
Ahmad Ali and Johra Khan	
3. Tools Used in Genetic Engineering	45
Ankita Sharma, Ahmad Ali, and Johra Khan	
4. Introduction of Recombinant DNA into Host Cells	77
Rida Saleem, Sitara Nasar, and Saima Iftikhar	
5. Linking of Desired Gene with DNA Vector/Gene Cloning Vector	97
Sitara Nasar and Saima Iftikhar	
6. Polymerase Chain Reaction	119
Asima Tayyeb and Zhuha Basit	
7. Concept and Nature of Genes	147
Anjum Sabba and Najeebul Tarfeen	
8. Blotting Techniques	169
Asima Tayyeb, Zhuha Basit, and Hanfa Ashfaq	
9. Chromosome Jumping	189
Muhammad Ishtiaq, Mahnoor Muzammil, and Mehwish Maqbool	
10. Electrophoresis	199
Muhammad Ishtiaq, Mahnoor Muzammil, and Mehwish Maqbool	

11. Genetically Engineered Microorganisms.....	219
Krina Mehta, Rohit Patel, Sameera Sharma, Arpit Shukla, Dweipayan Goswami, Meenu Saraf, and Paritosh Parmar	
12. Molecular Markers and Their Applications.....	251
Abdul Rehman, Hafiza Iqra Almas, Abdul Qayyum, Hongge Li, Zhen Peng, Guangyong Qin, Yinhua Jia, Zhaoe Pan, Shoupu He, and Xiongming Du	
<i>Index</i>.....	285

Contributors

Ahmad Ali

Department of Life Sciences, University of Mumbai, Vidyanagari, Santacruz (E), Mumbai, Maharashtra, India

Hafiza Iqra Almas

Department of Botany, University of Agriculture, Faisalabad, Pakistan

Hanfa Ashfaq

School of Biological Sciences, University of the Punjab, Lahore, Pakistan

Zhuha Basit

School of Biological Sciences, University of the Punjab, Lahore, Pakistan

Xiongming Du

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University, Zhengzhou, China

Dweipayan Goswami

Department of Microbiology and Biotechnology, Gujarat University, Ahmedabad, Gujarat, India

Shoupu He

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University, Zhengzhou, China

Saima Iftikhar

School of Biological Sciences, University of the Punjab, Lahore, Pakistan

Muhammad Ishtiaq

Department of Botany, Mirpur University of Science and Technology (MUST), Mirpur, Azad Jammu and Kashmir (AJK), Pakistan

Yinhua Jia

State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, Henan, China

Johra Khan

Department of Medical Laboratory Sciences, College of Applied Medical Sciences, Majmaah University, Majmaah, Saudi Arabia

Hongge Li

State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, Henan, China

Mohammad Amin Lone

Department of Zoology, Government Degree College, Uri, Jammu and Kashmir, India

Mehwish Maqbool

Department of Botany, Mirpur University of Science and Technology (MUST), Mirpur, Azad Jammu and Kashmir (AJK), Pakistan

Krina Mehta

Department of Microbiology and Biotechnology, Gujarat University, Ahmedabad, Gujarat, India

Mahnoor Muzammil

Department of Botany, Mirpur University of Science and Technology (MUST), Mirpur, Azad Jammu and Kashmir (AJK), Pakistan

Najeebul Tarfeen

Center for Research and Development (CORD), University of Kashmir, Hazratbal, Srinagar, Jammu and Kashmir, India

Sitara Nasar

School of Biological Sciences, University of the Punjab, Lahore, Pakistan

Zhaoc Pan

State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, Henan, China

Paritosh Parmar

Department of Biotechnology and Bioengineering, Institute of Advanced Research, Koba Institutional Area, Gandhinagar, Gujarat, India

Rohit Patel

Department of Microbiology and Biotechnology, Gujarat University, Ahmedabad, Gujarat, India

Zhen Peng

State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Science, Anyang, Henan, China

Abdul Qayyum

Department of Plant Breeding and Genetics, Bahauddin Zakariya University, Multan, Pakistan

Guangyong Qin

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University, Zhengzhou, China

Abdul Rehman

Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University, Zhengzhou, China

Rida Saleem

School of Biological Sciences, University of the Punjab, Lahore, Pakistan

Anjum Sabba

Department of Biochemistry, University of Kashmir, Hazratbal, Srinagar, Jammu and Kashmir, India

Meenu Saraf

Department of Microbiology and Biotechnology, Gujarat University, Ahmedabad, Gujarat, India

Anzar A. Shah

Department of Zoology, Government Degree College, Uri, Jammu and Kashmir, India

Ankita Sharma

Department of Life Sciences, University of Mumbai, Vidyanaigari, Santacruz (E), Mumbai, Maharashtra, India

Sameera Sharma

Department of Microbiology and Biotechnology, Gujarat University, Ahmedabad, Gujarat, India

Arpit Shukla

Department of Biological Sciences and Biotechnology, Institute of Advanced Research,
University of Innovation, Koba Institutional Area, Gandhinagar, Gujarat, India

Asima Tayyeb

School of Biological Sciences, University of the Punjab, Lahore, Pakistan



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Abbreviations

A	adenine
ADA	adenosine deaminase
AFLP	amplified fragment length polymorphism
Ald	α -acetolactate decarboxylase
ALP	alkaline phosphatase
AMV	avian myeloblastosis virus
ARMS	amplification refractory mutation system
ARS	autonomously replicating sequence
BACs	bacterial artificial chromosomes
BSA	bovine serum albumin
Bst	<i>Bacillus stearothermophilus</i>
C	cytosine
CarE B1	carboxylesterase B1
CCCP	carbonyl cyanide m-chlorophenyl hydrazone
cDNA	complementary DNA
CE	capillary electrophoresis
CRISPR	clustered regularly interspaced short palindromic repeats
CSF	cerebrospinal fluids
CVS	chorionic villus sampling
DB 71	direct blue 71
DEAE-dextran	diethylaminoethyl-dextran
DGGE	denaturing gradient gel electrophoresis
D-LDH	d-lactate dehydrogenase
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid

DNase I	deoxyribonuclease I
dNTPs	deoxyribonucleoside triphosphates
DOPE	dioleoylphosphatidyl ethanolamine
Dr	diacetyl reductase
dsDNA	double-stranded DNA
E. coli	Escherichia coli
EDTA	ethylene-diamine-tetraacetate
ELOSA	enzyme-linked oligonucleotide sorbent assay
EMF	electromotive force
ER	endonucleases
FBrAtio	flux balance analysis with flux ratios
FDA	Food and Drug Administration
FO	function oxidases
FRET	fluorescence resonance energy transfer
G	guanine
GE	genetically engineered
GEM	genetically engineered microorganisms
GEMs	genetically engineered microbes
GLP-1	glucagon-like peptide 1
GM	genetically modified
GMO	genetically modified organisms
GRAS	generally recognized as safe
GST	glutathione S-transferases
GTS	genomic template stability
HBV	hepatitis B virus
HFCS	high-fructose corn syrup
HFD	high-fat diet
HGT	horizontal gene transfer
HPCE	high-performance capillary electrophoresis

HPLC	high-performance liquid chromatography
HSV	herpes simplex virus
HTML-PCR	homopolymer tail-mediated ligation PCR
IEF	isoelectric focusing
IPTG	isopropyl β -d-1-thiogalactopyranoside
ITP	isotachophoresis
LAB	lactic acid bacteria
Ldh	lactate dehydrogenase
LIF	laser-induced fluorescence
lncRNA	long non-coding RNA
LTRs	long terminal repeats
MAB	marker-assisted breeding
MAC	mammalian artificial chromosome
MCE	microchip electrophoresis
MCF	microbial cell factories
MCS	multiple cloning site
MFO	mixed function oxidases
mpd	methyl parathion hydrolase gene
MT	metallothionein
OP	organophosphorus
ORF	open reading frames
Ori	origin of replication
PACs	P1-derived artificial chromosomes
PAGE	polyacrylamide gel electrophoresis
PAR α	proliferator-activated receptors α
PCR	polymerase chain reaction
PFE	pulsed-field gel electrophoresis
pHSA	plasma HSA
Pnk	polynucleotide kinase

PNP	p-nitrophenol
PPAR α	proliferator-activated receptors α
PTM	post-translational modifications
PVDF	polyvinylidene fluoride
pytH	pyrethroid hydrolase gene
QS	quorum sensing
rDNA	recombinant DNA
RFLP	restriction fragment length polymorphism
rHSA	recombinant HSA
RM	restriction/modification
RNA	ribonucleic acid
RT	real-time
RT-PCR	reverse transcriptase PCR
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SNP	single nucleotide polymorphism
sRNA	small regulatory RNAs
SSC	saline-sodium citrate
ST	S-transferases
T	thymine
TAE	tris/acetate/EDTA
YAC	yeast artificial chromosome

Acknowledgment

The editors extend their appreciation to the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia for supporting this work through Project No. GRANT3066.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Foreword

Genetic engineering is the science of manipulating the genetic material of an organism. The first artificial genetic modification accomplished by Herbert Boyer and Stanley Cohen in 1973 was the result of a series of advancements in techniques that allowed the direct modification of the genome by transferring genetic information from one organism to the other. Important advances included the discovery of restriction enzymes by Werner Arber and Hamilton O. Smith in the 1970s and DNA ligases by combined efforts of Gellert, Lehman, Richardson, and Jerard Hurwitz in the 1960s and early 1970s, the ability to design plasmids, and technologies like polymerase chain reaction (PCR) by Kary Mullis in 1983, and DNA sequencing by Frederick Sanger, Walter Gilbert, Allan Maxam in 1977 and John Craig Venter in 1999.

Several other discoveries and developments that occurred later brought this technology of genetic engineering to the level that we see today. From 1990 to 2003, the Human Genome Project succeeded in mapping the human genome with more than 20,000 genes identified and their genomic loci documented. Dolly was cloned under the leadership of Ian Wilmut in 1996. Transformation of the DNA into a host organism was accomplished with the invention of biolistics, agrobacterium-mediated recombination, and microinjection. The first genetically modified (GM) animal was a mouse created in 1974 by Rudolf Jaenisch. In 1976 the technology was commercialized, with the advent of genetically modified bacteria that produced somatostatin, followed by insulin in 1978. In 1983 an antibiotic-resistant gene was inserted into tobacco, leading to the first genetically engineered (GE) plant.

Advances followed that allowed scientists to manipulate and add genes to a variety of different organisms and induce a range of different effects. Plants were first commercialized with virus-resistant tobacco released in China in 1992. The first genetically modified food was the Flavr Savr tomato, marketed in 1994. By 2010, 29 countries had planted commercialized biotech crops. In 2000 a paper published in *Science* introduced golden rice, the first food developed with increased nutrient value.

In the present decade of 2010–2020, the development therapies that have been theorized and studied for years are finally being approved for humans, as the discovery of CRISPR, possibly the greatest achievement in the history of genetics, takes hold in the world.

Volume 1 of this book, *Genetic Engineering*, encompasses all the basic concepts of the various components of recombinant DNA (rDNA) technology. The book has been designed to communicate the fundamental principles of genetic engineering through an explanation of various molecular biology phenomena and the development of various technologies over the years. The various chapters have been written by young and experienced experts in their respective fields from different elite educational institutions across the globe. I applaud the editor, Dr. Tariq Ahmad Bhat, as well as the book chapter contributors for successfully bringing together this volume.

—**Abdul Rauf Shakoori**

Distinguished National Professor,

Professor Emeritus,

University of the Punjab, Lahore, Pakistan

Preface

The advent of biotechnology has forever changed human perception of living entities. Genetic engineering enables the precise control of the genetic composition and gene expression of organisms directed toward the advantage of human well-being. Innovative applications have emerged in environmental sustainability, food and nutritional security, and medicinal advancement. The utilization of this powerful technology solicits contrasting opinions among scientists, politicians, and the public in relation to biosafety and bioethics. This necessitated the engagement in research aimed at understanding the safety of genetic engineering products and prompted the development of national and international policies. This book addresses these aspects in two volumes: *Volume 1, Genetic Engineering: Principles, Mechanism, and Expression*, and *Volume 2, Genetic Engineering: Applications, Bioethics, and Biosafety*.

Volume 1 consists of 12 chapters covering genetic engineering concepts, molecular tools, and technologies utilized in the manipulation, amplification, and introgression of DNA. Topics covered are concepts of genetic engineering, enzymes of genetic engineering, tools used in genetic engineering, the introduction of recombinant DNA (rDNA) into host cells, linking of the desired gene with DNA vector/gene cloning vector, polymerase chain reaction (PCR), concept and nature of genes, blotting techniques, chromosome jumping, electrophoresis, genetically engineered (GE) microorganisms, and molecular markers and their applications.

This book is a valuable asset to upper-undergraduate and postgraduate students, teachers, and researchers interested in cell biology, genetics, molecular genetics, biochemistry, biotechnology, botany, zoology, and agriculture sciences. The chapters are contributed by experts in their fields, presenting recent contemporary developments in genetic engineering research supported with illustrations, tables, and recent references. We are thankful to all the authors across the globe who contributed their research output in the form of book chapters to make our project a successful endeavor. We wish to present our gratitude to the contributing authors for their generous cooperation and to Apple Academic Press (AAP)/CRC

Press/Taylor and Francis Group for giving us the opportunity to publish this work.

—**Tariq Ahmad Bhat, PhD**

Jammu and Kashmir, India

Jameel M. Al-Khayri, PhD

Al-Ahsa, Saudi Arabia

CHAPTER 1

Concepts of Genetic Engineering

MOHAMMAD AMIN LONE and ANZAR A. SHAH

Department of Zoology, Government Degree College, Uri, Jammu and Kashmir, India

ABSTRACT

Genetic engineering, also known as gene modification or gene editing, is a field of biotechnology that involves the direct manipulation of an organism's genetic material in order to modify or add traits. This can be done through a variety of techniques, such as the insertion of genetically modified DNA into an organism, or the disruption or suppression of certain genes. Genetic engineering has the potential to revolutionize the way we produce food, create new medicines, and address environmental challenges. However, it is also a controversial field, with many ethical, social, and environmental considerations.

One of the key concepts in genetic engineering is the use of genetically modified organisms (GMOs). These are organisms whose genetic material has been modified using genetic engineering techniques. The goal of using GMOs is often to introduce new traits or characteristics into the organism, such as increased resistance to pests or diseases, or enhanced nutritional value. However, there are concerns about the safety of GMOs, both for the environment and for human consumption.

Another important concept in genetic engineering is gene editing, which involves the precise modification of an organism's genetic material at the DNA level. Gene editing has the potential to be used for a wide

range of applications, including the treatment of genetic diseases and the production of new medicines. However, there are also concerns about the potential ethical implications of gene editing, such as the creation of designer babies or the enhancement of human traits.

Overall, genetic engineering is a complex and rapidly-evolving field with the potential to significantly impact many aspects of our lives. While it offers many benefits, it also raises important ethical, social, and environmental considerations that need to be carefully considered as the field continues to develop.

1.1 HISTORICAL PERCEPTIVE OF GENETIC ENGINEERING

The structural elucidation of genetic material and code marked a new dawn in the field of science. It became very evident that genetic material conceals numerous biological secrets inside, which were yet to be explored. Post-1970s led to huge technological advancement, which paved the way towards new analysis and manipulation in genetic material. Consequently, many breakthroughs were achieved after the detailed structure of DNA and RNA was documented. It was only after World War II, people mainly biologists and philosophers began to analyze a growing link between biology and technology. Eventually, a new field of science which could be used for the improvisation of human life evolved as genetic engineering. The idea of changing the human nature may thus benefit the society. It was only after Second World War, scientists began to institutionalize biology and technology by establishing new departments, institutes, and ministries. Thus, with the new advances in the field of molecular biology, genetic engineering became more of science than merely an art.

The concept of genetic engineering dates long before the structural documentation of DNA and RNA. However, the expression “genetic engineering” was first used by the science-fiction novelist but not by scientist. It was in 1951; Jack Williamson used the term “genetic engineering” in his novel “Dragon’s Island” (Williamson, 2012).

People around the globe were already familiar with the art of using living micro-organisms to manufacture new products. The best example in this regard is the process of fermentation, where ancient people utilized the service of living micro-organisms to get new and improved products such as bread and alcohol. As the events progressed, knowledge of protein

synthesis from regions of deoxyribonucleic acid (DNA) (called genes) provided a key to in-depth exploration in the DNA structural base.

With plenty of exploration and subsequent advancement of new genetic techniques, expression of gene was made conceivable regardless of its origin, in a simple micro-organism like *Escherichia coli* (*E. coli*) leading to the increased production of the product coded by that gene. Based on the same principle, vast number of other organisms both micro and macro were used to get efficient and large quantity of products. It was only before 1990 a new technique namely protein engineering became possible as a new outcome of genetic engineering (Maloy & Hughes, 2013). With the course of time, new discoveries and advancements came into existence which changed the whole scenario of the scientific world. It was Karl Ereky who specified a whole new subject as biotechnology (Fári & Kralovánszky, 2006). It is very important to mention that the term genetic engineering was coined in 1940 by A Jost. However, the most important discovery in the arena was that of enzymes and vectors. These two together changed the whole concept of biology and hence marked the beginning of a new dawn in the scientific world.

1.2 KEY CONCEPTS IN GENETIC ENGINEERING

1.2.1 GENE

Gene is the basic physical, structural, and functional unit of inheritance which occupies a fixed place on a chromosome. It may be chemically defined as a sequence of nucleotides carrying specific hereditary information (Pearson, 2006). More specifically, it is the nucleic acid DNA or RNA which codes the synthesis of the gene product, whether it is RNA or protein (Morange, 2000).

1.2.2 RECOMBINANT DNA (RDNA)

Recombinant DNA (rDNA) can be defined as a solitary chimeric DNA produced by joining two or more dissimilar fragments of DNA from different organisms. The technique used to produce rDNA is called rDNA technology. The organism is categorized as donor and host. The donor is the organism from which the isolation of DNA is done and the host is the

organism in which this DNA is incorporated. Paul Berg, Herbert W Boyer, and Stanley N Cohen are the pioneers of rDNA technology. The first ever rDNA molecule was created by recombining the SV40 monkey DNA virus genome and a bacterial virus known as phage λ (Robl, Wang, Kasinathan, & Kuroiwa, 2007).

1.2.3 VECTORS-DELIVERY SYSTEM IN RECOMBINANT DNA (RDNA) TECHNOLOGY

Vectors are important tool in rDNA technology as they act as transporters and thereby work as vehicles which deliver the genetic material in the organism of interest. Vectors may thus be defined autonomously replicating DNA molecule engaged to transport foreign hereditary material to host cells. Vectors are basically transgenic DNA bodies which possess a larger unit as foundation and smaller unit as foreign DNA. The smaller foreign DNA is ultimately expressed in the host cell thereby is transmitted. Vectors are broadly classified into two types, i.e., cloning vector and expression vector. Cloning vector is used to amplify the number of copies of a cloned DNA fragment. However, Expression vector is designed to for expressing foreign gene into a protein.

1.3 PROPERTIES OF AN EFFICIENT AND IDEAL VECTOR

An efficient vector should possess the following properties:

- The first and foremost property which an ideal vector should possess is an autonomous replicating nature which means it should possess Ori (origin of replication) region.
- It should possess at least a selectable marker, e.g., antibiotic resistance marker.
- Should possess a screenable marker/scorable marker. Scorable marker produces an end product which can easily be noticed by means of a simple and quantitative assay. The examples include markers of β -galactosidase, green fluorescent protein, etc.
- Occurrence of unique restriction enzyme site. This is very important as it forms the main criterion for any vector to be used in rDNA technology.

- Should possess several cloning sites.
- Should possibly be small in size and easy to handle.
- Should kick off DNA replication autonomously to get numerous copies.
- Presence of suitable regulatory elements for expressing foreign gene.
- The selection of a suitable vector depends mainly on the size limit of inserted DNA and category of host projected for cloning.

The other classification of vectors is based on its transmitting form and includes synthetic or artificial chromosomes, plasmids, viruses, and cosmids. However, of all four types, the most commonly used vectors include viruses and plasmids. Vectors have the property of being transcribed, translated, inserted, and then expressed.

1.4 PLASMIDS

These vectors include double-stranded spherical DNA having the property of self-replication in the host. They are primarily characteristic of prokaryotes but can also exist in eukaryotes. Because of autonomous self-replicating behavior, they are also known as replicons. Plasmids mainly carry genes for antibiotic resistance which later on acts as marker in rDNA technology. Plasmids are of two types, i.e., integrating plasmids and non-integrating plasmids. As the name specifies, non-integrating plasmids don't unite with the host DNA and replicate autonomously when the host cell divides. However, integrating plasmids amalgamate with the host DNA to get replicated. The size of plasmid ranges from few kbs to 100 kbs. Examples of plasmid vectors include pBR322 and pUC vector (Figure 1.1).

1.5 VIRUSES

Viruses play a key role as carriers and execute this role by *in vivo* and *in vitro* methods. Viruses transfer genes by the process of transduction. Viral vectors are primarily pathogenic but their pathogenicity is minimized by remodeling their genetic structure and by scissoring the gene of replication. These viruses can only infect and can't reproduce rendering them

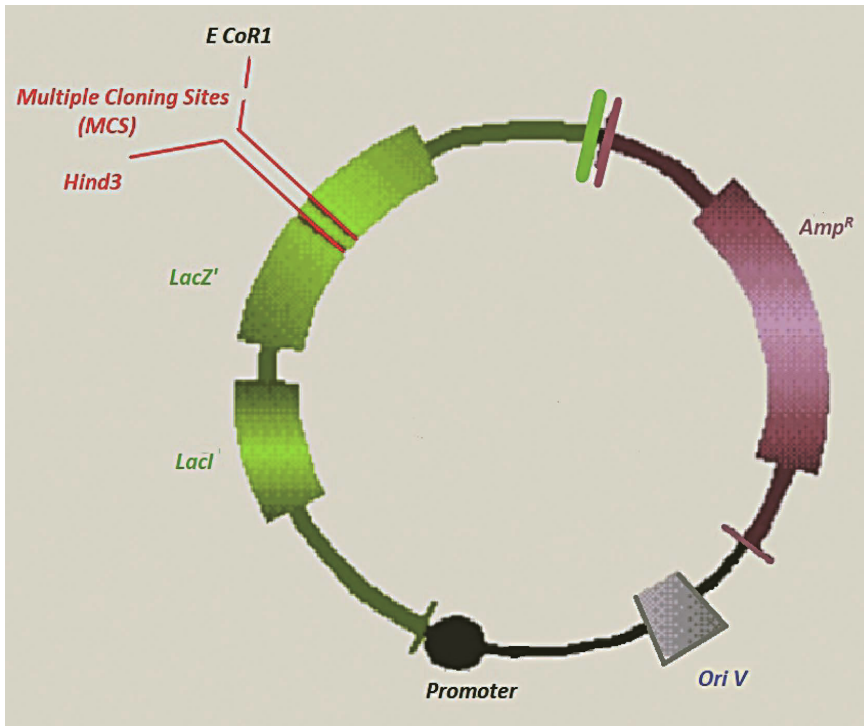


FIGURE 1.1 Diagram showing a plasmid vector construct from the pUC family consisting of ampicillin-resistant gene, promoter, LacI repressor gene, origin of replication LacZ gene for peptide of galactosidase. It also contains multiple cloning regions showing different restriction endonucleases.

harmless. Viruses are very stable and concrete. Viral vectors can infect multiple types of cells thereby are diverse in nature compared to plasmid. Avery, a smaller number of naturally occurring viruses manufacture single-stranded DNA in their life cycle. Examples include M13, fl, and fd and have been manipulated to be suitable vector (Figure 1.2).

1.6 COSMID

The huge drawback of the plasmid vector is its transformation inefficiency of cloning large segments of DNA. In order to overcome this problem, a new type of vector has been designed with high effectiveness compared to the plasmid vector. The cosmid vector is the blend of the plasmid

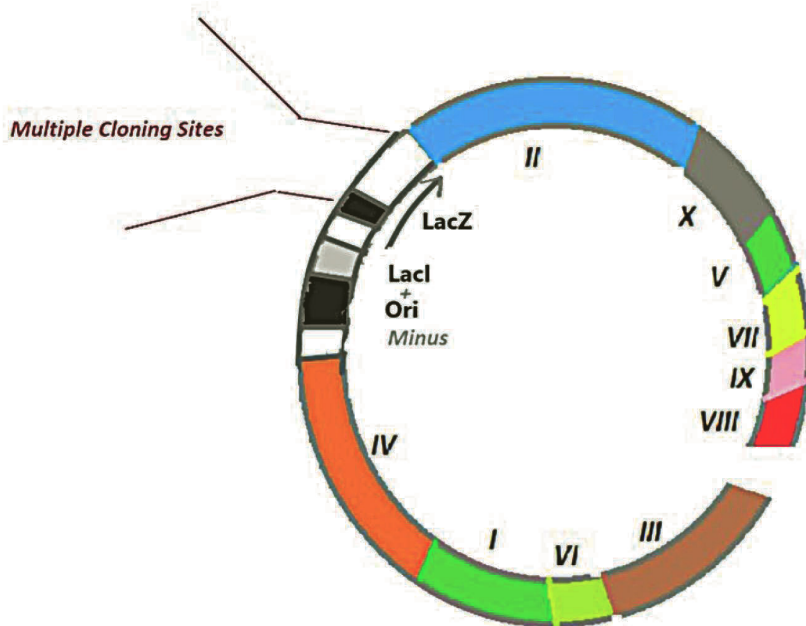


FIGURE 1.2 Diagram showing M13mp18 viral vector showing multiple cloning sites.

vector and the lambda phage *cos* sequence (Hohn & Murray, 1977). These arrangements enable target DNA to be incorporated into the λ head (Chauthaiwale, Therwath, & Deshpande, 1992). They were first described by Collins and Hohn in 1978.

The cosmid vector can hold up to 45 kb of DNA while plasmid and λ phage vectors are limited to 25 kb (Hohn & Murray, 1977).

The main difference compared to usual plasmid vectors is the presence of a small piece of lambda DNA, known as the cohesive end site. Lambda DNA in the virus is linear with two complementary single-stranded ends, therefore, are called the cohesive ends or *cos* sites (Figure 1.3).

1.7 SYNTHETIC OR ARTIFICIAL CHROMOSOMES

These are also cloning vectors that transport larger DNA inserts than plasmids or lambda-phage-derived vectors. Artificial chromosomes contain all

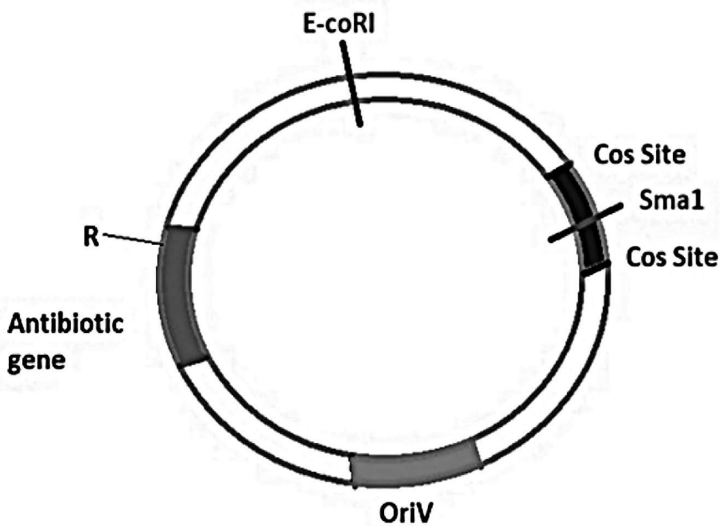


FIGURE 1.3 Diagram showing structure of cosmid vector–plasmid vector + Cos site.

the elements that are critical for replication and stability of the molecule within the host cell (O’Brien & Lummis, 2011). For instance, the first artificial chromosome, known as yeast artificial chromosome (YAC) vectors, is able to bear DNA inserts as huge as 2,000 kb. Similarly, bacterial artificial chromosomes (BACs) and P1-derived Artificial chromosomes (PACs) were designed to address the difficulty of insert chimerism and instability (Robl et al., 2003). Other examples include mammalian artificial chromosomes (MACs) designed for mammalian cells, including human cells (Figure 1.4 and Table 1.1).

1.8 BASIC STEPS REQUIRED IN GENETIC ENGINEERING

1.8.1 ISOLATION OF DNA

Genetic material is contained inside the cells. It has to be obtained in pure form without even the attached histones and other proteins. The first step in DNA purification is to open the cells and release DNA. The method should be gentle to preserve the native DNA; the approaches to break the cells vary due to the variability in the cell structure.

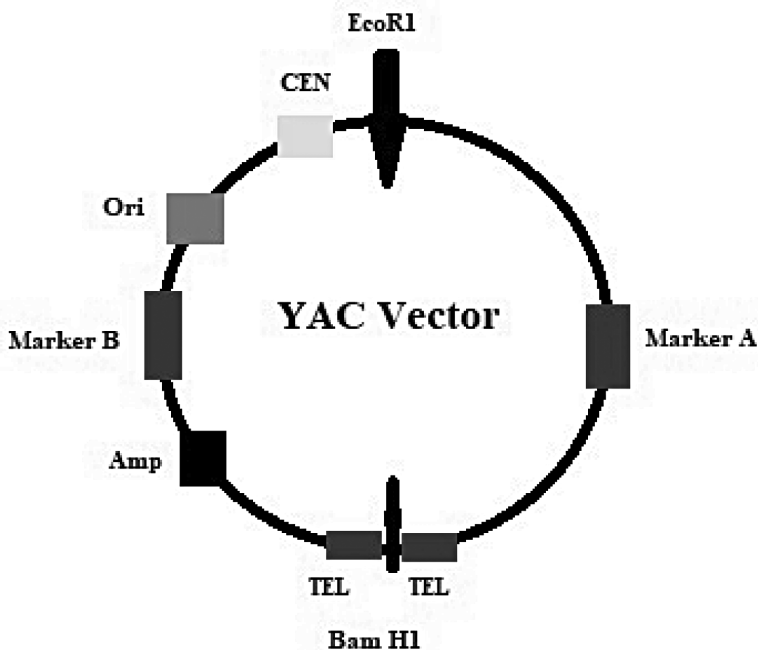


FIGURE 1.4 Diagram showing YAC vector.

TABLE 1.1 Some Milestones in the History of Genetic Engineering

Sl. No.	Year	Event
1.	1917	The term Biotechnology coined by Karl Ereky.
2.	1940	The term Genetic Engineering coined by A. Jost.
3.	1943	Penicillin produced on an industrial scale.
4.	1944	Avery, Macleod, and Mc Carty demonstrated that DNA is the genetic material.
5.	1953	The structure of DNA was determined Watson and Crick.
6.	1958	For studying the primary structure of proteins and genetic recombination in a bacterium F. Sanger, Joshua Lederberg, G.W. Beadle, and Edward L. Tatum were jointly awarded the Noble prize in chemistry and physiology or Medicine.
7.	1959	Arthur Kornberg and Severo Ochoa were awarded the Noble prize for the synthesis of DNA and RNA.
8.	1961–1966	The entire genetic code is deciphered.
9.	1970	The First Restriction endonuclease, Hind II was isolated by Hamilton O. Smith, Thomas Kelly, and Kent Wilcox from the bacterium <i>Haemophilus influenza</i> .

TABLE 1.1 (Continued)

Sl. No.	Year	Event
10.	1973	Boyer and Cohen established recombinant DNA technology.
11.	1974	Rudolf Jaenisch created a genetically modified mouse.
12.	1977	Fredrich sanger developed a method for sequencing DNA.
13.	1978	Genentech produced human insulin in <i>E. coli</i>
14.	1983	PCR developed by Karry Mullis.
15.	1983	An antibiotic-resistant gene was introduced into tobacco, leading to the first genetically modified plant.
16.	1987	Recombinant vaccine for hepatitis B (HBsAg) became the first synthetic vaccine for public use.
17.	1988	PCR method was published.
18.	1990	The human genome was officially initiated.
19.	1994	The first GM food known as FlavrSavr tomato has been commercialized globally. The FDA reports that genetically modified tomatoes are as safe as traditional tomatoes.
20.	1995	The first genome sequence for the bacterium <i>Haemophilus influenza</i> has been completed.
21.	1996	The commercial planting of genetically modified crops has been launched.
22.	1997	India became the fourth country in the world to develop an indigenous hepatitis B vaccine.
23.	1997	The first-ever mammal clone named Dolly was developed by Wilmut and Campbel. Dolly was a sheep with a mother and no father.
24.	2000	GM rice, called golden rice, rich in beta carotene, the precursor of Vitamin A, has been developed.
25.	2001	Human genome is sequenced.

1.8.1.1 LYSIS OF CELLS

The lysis or breakdown of bacterial cells is brought about by the use of an enzyme Lysozyme and the chemical ethylene-diamine-tetraacetate (EDTA) followed by the addition of detergents, namely sodium dodecyl sulfate (SDS).

The lysis of animal cells is directly carried out by treating them with a detergent-like SDS. Plant cells require very harsh treatment to break them

open. Besides the treatment of enzymes like *cellulase* and *pectinase*, the cells are frozen and subsequently crushed in mortar and pestle.

1.8.1.2 PURIFICATION OF DNA

This involves the complete breakdown or removal of all the cellular materials other than DNA. This is achieved by homogenizing and centrifuging the treated cells to break the cells and their nuclear envelopes. The homogenized product is then treated with proteases to digest histones and other proteins; ribonucleases to digest RNA; amylases and lipases to digest polysaccharides and lipids, respectively. DNA remains intact. This DNA is precipitated by adding ethanol. DNA appears as a mass of very fine threads that can be separated by spooling or winding over a fine appliance.

1.8.2 GENERATION OF DNA FRAGMENTS

The desired genes (gene of interest) to be cloned are prepared in three ways:

1. In the first method, the gene of interest is isolated from the total genomic DNA of an organism. To isolate a gene, the genomic DNA is treated with a restriction endonuclease enzyme to cut it into many chunks. These pieces/fragments of DNA are then separated according to their length by electrophoresis. The separated DNA segments are finally used for cloning.
2. In the second method, the desired gene is synthesized using reverse transcription of the mRNA of the gene. The new DNA formed by this process is known as complementary DNA (cDNA).
3. In the third method, the preferred gene is produced by a programmed machine referred to as a DNA synthesizer or gene machine. The gene is thereby amplified by PCR and desired DNA (gene of interest) is referred to as transient or target DNA or donor DNA.

1.9 ISOLATION OF DNA VECTOR

The DNA, which is used for transferring the desired DNA (gene of interest) into a host cell or organism, is known as a vector. It is also called a cloning vehicle or carrier molecule. They are self-replicating in an appropriate

host cell. Vectors are of two kinds: plasmids (e.g., pBR322) and DNA viruses (e.g., lambda phages).

1.10 CONSTRUCTION OF RECOMBINANT DNA (RDNA)

Both the passenger and vector DNAs are separated with the same restriction endonucleases (ER) to produce complementary sticky ends. Self-ligation is prevented by using *alkaline phosphatase* (ALP). Both passenger and vector DNAs are now joined together using DNA ligase (Gupta, Sengupta, Prakash, and Tripathy, 2017). The vector is guaranteed to receive only one passenger DNA. Inserting the selected DNA (passenger DNA) into the cloning vector (e.g., plasmid) produces rDNA. The rDNA is also called Chimeric DNA. Chimera in Greek Mythology refers to a monster that has a lion's head, a goat's body, and a serpent's tail.

1.11 INTRODUCTION OF RDNA INTO HOST CELL

Introducing rDNA into a host cell is a significant step in genetic engineering. It is the efficiency of this process that determines success. This is accomplished by any of the methods discussed in subsections.

1.11.1 TRANSFORMATION

Direct transformation of DNA fragments in the medium by bacterial cells is called transformation. This method was first adopted in 1970 by Mandell and Higa (1970). The ability of bacterial cells to take DNA from the medium is known as competence.

In transformation, the alien DNA is injected into a bacterial cell (for example, E-COLI). The uptake of Foreign DNA by E-coli is carried out in ice-cold CaCl_2 (0–5°C) trailed by heat shock (37–45°C) for about 90 seconds.

1.11.2 TRANSFECTION

The direct intake of naked DNA from the culture medium by eukaryotic cells is termed as transfection (Neumann, Schaefer-Ridder, Wang, & Hofschneider, 1982). Transfection is similar to transformation in the case of bacteria. It occurs in plant cells and animal cells.

In this method, DNA is first dissolved in the phosphate buffer and Calcium chloride solution is added to it which leads to the formation of calcium phosphate precipitate. The precipitate is then added to the cells which lead to the adherence of the precipitate particles with the cell surface. The cells engulf the particles along with DNA by the process of phagocytosis. The DNA entering the cell is integrated with the cell's genome.

1.11.3 CONJUGATION

A naturally occurring microbial recombination also referred to as bacterial mating which occurs when two bacteria (one donor and one recipient) meet together, join by a cytoplasmic bridge or conjugation tube, and exchange the DNA (from donor to recipient). Within the recipient cell, the new DNA either integrates with the chromosome or remains free.

1.11.4 ELECTROPORATION

Electroporation works on the principle of high voltage electrical impulses. Electroporation is a technique that uses membrane permeability through the electric field. Neumann et al. reported in 1982, how alien DNA can be introduced into mouse cells using short pulses of high voltage electric field (Neumann et al., 1982). It causes the absorption of DNAs into protoplasts through the temporary permeability of the plasma membrane to macromolecules. Electric shocks generate pores in the membrane by damaging them. DNA diffuses through these membranes immediately after the electric field is applied. Transformation of *E. coli* cells by electroporation was finally published in 1988 (Carter & Shieh, 2015).

1.11.5 GENE TRANSFER THROUGH LIPOSOMES

Liposomes are circular lipids with an aqueous interior that can transport nucleic acids. The liposome-mediated gene transfer is referred to as Lipofection (Carter & Shieh, 2015). This technique was used as a carrier for the incorporation of nucleic acids into plant protoplasts. The liposome-treated pieces of DNA are encapsulated in them. These liposomes may then adhere

to cell membranes and merge with them to transfer DNA fragments. In this way, DNA penetrates the cell and then into the nucleus.

1.11.6 DIRECT TRANSFER OF DNA

1.11.6.1 PARTICLE BOMBARDMENT OR BIOLISTICS

Shooting of plant or animal cells by DNA-coated gold or tungsten particles for introducing DNAs into cells. It is also known as microprojectile bombardment. This method is employed for incorporating rDNA into plant cells, fungal cells, animal cells, and cell organelles such as mitochondria and chloroplast (O'Brien & Lummis, 2011). The instrument used to shoot the DNA into the cells is called a Gene gun or Microprojectile gun. Vector is not used in this method to incorporate rDNA into the host cells. The gene gun is used in place of Vector.

1.11.6.2 MICRO-INJECTION

It refers to the injection of DNAs or cell organelles directly into host cells using an injection method. This DNA is usually employed to inject the DNA directly into the animal cells, eggs, zygotes, and plant protoplasts. In this method, fertilized egg (Host cell) is transferred into a microscopic slide placed under the microscope. The host cell is then maintained in place using a suction pipette. One end of the sucking pipette is positioned on the surface of the cell and gentle suction is applied on its other end. The rDNA is sucked into a glass injection needle and is gently inserted into the zygote by viewing through the microscope. The rDNA is delivered into the zygote and the needle is drawn back carefully. In this way, the rDNA gets into the genome of the zygote.

1.12 SELECTION AND MULTIPLICATION OF RECOMBINANT HOST CELLS

Selection is a fundamental process in which, when the rDNA is inserted into a particular host cell, it becomes necessary to detect the cells that have received the recombinant or foreign DNA molecule. This process

is referred to as Screening or Selection. A Host cell containing rDNA is called recombinant or transformant. In this step, the transformed cells are separated from the non-transformed cells by using various methods such as antibiotic resistance, colony hybridization, and blotting tests, etc., using marker genes. The recombinant cells are mass cultured to test the purity of the products of the cloned genes.

1.13 EXPRESSION OF THE CLONED GENES

Finally, consideration should be given to ensure that the gene of interest is expressed as a functional protein. It is isolated and immunologically tested. The recombinant producing pure product (e.g., insulin) can readily be used in the future. The transformed host cells are also multiplied to produce enough copies (Figure 1.5).

1.14 APPLICATIONS OF GENETIC ENGINEERING

Manipulation of genetic constitution of organisms by inserting a gene of interest is called Genetic engineering. In genetic engineering, a novel gene can be incorporated into closely linked organisms as well as unrelated organisms. Through genetic engineering, many novel strains of micro-organisms, plants, and animals have been developed. These are very useful in agriculture, industries, pharmaceuticals, and environment.

1.14.1 TRANSGENIC ANIMALS

Transgenic animals are developed by manipulating their genetic makeup and introducing a new gene which is later expressed. These animals are utilized in the following ways:

- Transgenic animals are purposely designed to understand the control of genes and their impact on the normal functioning of the body and its developments. The best example in this regard is studying the biological role of insulin-like growth factor.
- The transgenic animals are developed to boost our understanding of result of genes which have role in development of diseases. Eventually, transgenic animals serve as model for human diseases.
- Transgenic animals are developed by introduction of a portion of DNA that encodes a product from the organism. Some of the

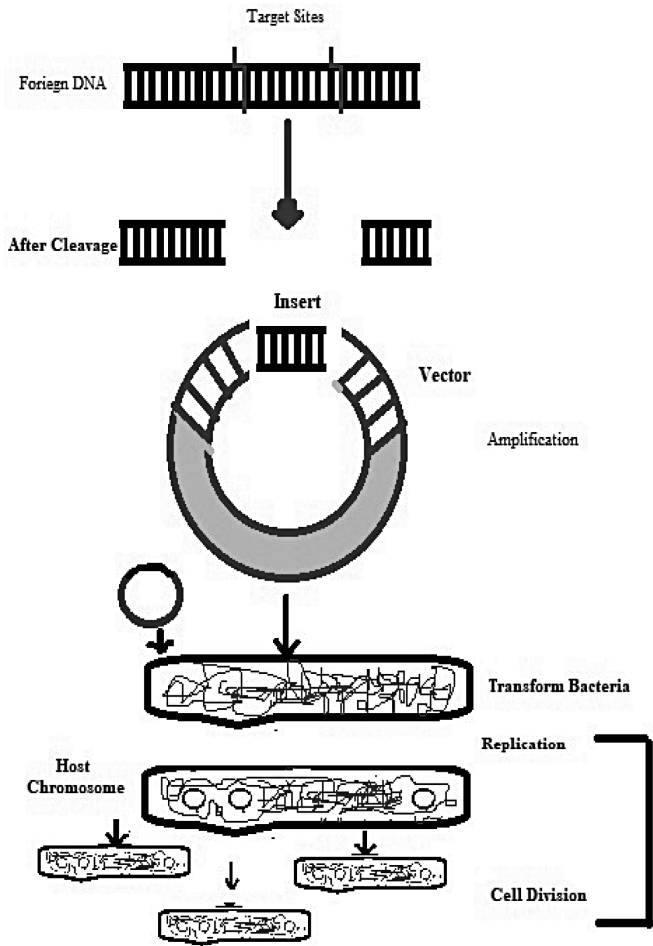


FIGURE 1.5 The process of genetic engineering.

useful biological compounds produced by this include a human protein namely alpha-1 antitrypsin used to treat emphysema. The first transgenic cow, Rosie, formed the human protein-rich milk; besides it possessed human alpha-lactalbumin, a nutritionally balanced product for babies. Transgenic cows with K-casein gene-insert produce milk rich in K-casein. This milk is likely for cheese making. Transgenic mice are also produced to test the efficacy and safety of vaccines like COVID-19 and HIV.

- Transgenic animals with high sensitivity to toxins are produced to observe and analyze the toxicity of drugs.
- Transgenic goats with tPA gene produce plasminogen activator in the milk. The tPA is used to dissolve blood clots.

1.14.2 TRANSGENIC PLANTS

The exercise of genetically modified (GM) plants has revolutionized the whole world and are very handy in the following ways:

- Genetic modification resulted in productions of better crops which are very resistant to abiotic stresses like frost, high temperature, drought, and high osmolarity, etc. (Davies, 2007).
- The use of synthetic pesticides has been declined as the crops are originally being GM for pest resistance.
- Damage to the crop after harvesting is also reduced.
- The nutrient content of soil is also preserved due to efficient mineral uptake by these crops.
- Food produced from GM crops has better dietary and nutritional value.
- Genetic alteration has been employed to generate need specific plants to deliver resources to industries like starch, fuel, pharmaceuticals, etc.

1.14.3 PRODUCTION OF PEST RESISTANT PLANTS

1.14.3.1 BT COTTON

The soil bacterium *Bacillus thuringiensis* transcribes crystal proteins known as *cry proteins*. These proteins are lethal to larvae of insects like Tobacco budworm, beetle worms and mosquitoes. The cry proteins are present in inactive form as protoxins and get transformed into active form when internalized by the organism, as the alkaline pH of gut solubilizes the toxin. The active toxin gets attached to the exterior of epithelial cells of midgut and generates pores. This results in enlargement and lysis of cells which ultimately causes the death of the Larva. These genes (cry genes) are isolated from the bacterium and integrated into numerous plants like tomato, cotton, corn, rice, soybean, etc. (Newell, 2000).

Below are the names of some of the cry genes encoding proteins against pests:

- Cry I Ac and cry II Ab manages cotton bollworms;
- Cry I Ab manages corn borer;
- Cry III Ab manages Colorado potato beetles.
- Cry III Bb manages corn rootworm.

1.14.3.2 DEFENSE AGAINST NEMATODES

Yield of tobacco plant is greatly affected by nematode namely *Meloidogyne incognita*. In order to prevent this. A special gene from the parasite is incorporated into the plant by means of *Agrobacterium* vector. The genes are incorporated in such a way that both sense and antisense RNA get formed. Being complementary to each other, they form double-stranded ribonucleic acid (RNA). As a result of this, specific RNA of nematode gets neutralized. This process is called RNA-interference. The whole process result nullifies the host and cannot assemble with a transgenic host which contributes in the host plant defense.

1.14.3.3 USE OF RDNA TECHNOLOGY IN THE MEDICAL FIELD

rDNA technology has paved the way in the production of highly efficient therapeutic drugs. Furthermore, therapeutics produced by recombination does not produce any harmful and unnecessary immunological response, which is very common in analogous products isolated from non-human resources. The current tally of approved recombinant therapeutics is 30. Out of these 30, 12 are already marketed in India.

1.14.3.4 HIMULIN – GENETICALLY ENGINEERED (GE) INSULIN

Himulin is genetically engineered (GE) form of insulin. Structurally human insulin contains two polypeptide chains as chain A and chain B, bonded by disulfide bridges.

Insulin is produced in an inactive form as prohormone and has to be processed before it attains complete functional structure. During its maturation, a polypeptide, namely C-peptide is scissored and maturation is attained. American company namely Eli-Lilly in 1983 synthesized two

DNA sequences responsible for producing chain A and B of functional insulin. They introduced this DNA in *E. coli* plasmid thereby producing insulin. Polypeptide chains produced were harvested and linked by the introduction of disulfide bonds.

1.14.3.5 GENE THERAPY

With the advancement of genetic engineering, it is now possible to treat certain genetic diseases. The method relays on the introduction of genes in cells and tissues of that individual. The defective gene is primed and new normal gene is delivered into the embryo or individual. Eventually the normal gene replaces the faulty mutant allele. The new genes are introduced in individual, and the process is mediated by special transporters called vectors. Viruses which infect the host and introduce their own genetic material there are designed in such a way that they act as vectors. The first documented gene therapy was done in the 1990s to a four-year-old girl, who had a deficiency of adenosine deaminase (ADA). Since ADA deficit can be corrected by bone marrow transplantation in certain children but it is not fully remedial. Lymphocytes are developed in a functional ADA. cDNA is incorporated into lymphocytes. These lymphocytes are then transferred in patient's body. A permanent cure is only possible if the same gene is incorporated in the patient's bone marrow cells in the embryonic stage.

1.14.3.6 MOLECULAR DIAGNOSIS

rDNA techniques like PCR (polymerase chain reaction) are useful in the early diagnosis of disorders. The cloned genes are also employed as 'probes' to spot the existence of complementary DNA strand. Eventually, a probe is a single-stranded DNA that is labeled with a radioactive tag and is used to discover its complementary DNA by hybridization. Afterwards, recognition of radioactivity by autoradiography is done. Existence of a normal or mutant gene can be identified by this method. Thus, PCR is a very powerful tool in the detection of genetic disorders besides other diseases like HIV. During the current global crises of COVID-19, PCR has served one the best and only technique to identify the presence of the virus.

1.15 SAFETY CONCERNS REGARDING GENETIC ENGINEERING

Genetically modified organisms (GMO) are primarily developed by introducing a minute segment of DNA from a “donor” into a “recipient” organism. The genome of the new organism is thus most likely that of the recipient organism. Consequently, the new organism contains fraction of recipient thus accessing the properties of recipient may provide as initial insight and assessment of properties of the new organism. Thus, a framework for safety assessment mainly depends on a range of differences in the genetic makeup of recipient and modified organism.

KEYWORDS

- **bacterial artificial chromosomes**
- **deoxyribonucleic acid**
- ***Escherichia coli***
- **mammalian artificial chromosomes**
- **P1-derived artificial chromosomes**
- **yeast artificial chromosome**

REFERENCES

- Carter, M., & Shieh, J. C., (2015). *Guide to Research Techniques in Neuroscience*. Academic Press.
- Chauthaiwale, V. M., Therwath, A., & Deshpande, V., (1992). Bacteriophage lambda as a cloning vector. *Microbiological Reviews*, 56(4), 577–591.
- Collins, J., & Hohn, B., (1978). Cosmids: A type of plasmid gene-cloning vector that is packageable *in vitro* in bacteriophage lambda heads. *Proceedings of the National Academy of Sciences*, 75(9), 4242–4246.
- Davies, K. M., (2007). Genetic modification of plant metabolism for human health benefits. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 622(1, 2), 122–137.
- Fári, M. G., & Kralovánszky, U. P., (2006). The founding father of biotechnology: Károly (Karl) ereky. *International Journal of Horticultural Science*, 12(1), 9–12.
- Gupta, V., Sengupta, M., Prakash, J., & Tripathy, B. C., (2017). *Basic and Applied Aspects of Biotechnology*. Springer.

- Hohn, B., & Murray, K., (1977). Packaging recombinant DNA molecules into bacteriophage particles *in vitro*. *Proceedings of the National Academy of Sciences*, 74(8), 3259–3263.
- Maloy, S., & Hughes, K., (2013). *Brenner's Encyclopedia of Genetics*. Academic Press.
- Mandel, M., & Higa, A., (1970). Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology*, 53(1), 159–162.
- Morange, M., (2000). *A History of Molecular Biology*. Harvard University Press.
- Neumann, E., Schaefer-Ridder, M., Wang, Y., & Hofschneider, P., (1982). Gene transfer into mouse lyoma cells by electroporation in high electric fields. *The EMBO Journal*, 1(7), 841–845.
- Newell, C. A., (2000). Plant transformation technology. *Molecular Biotechnology*, 16(1), 53–65.
- O'Brien, J. A., & Lummis, S. C., (2011). Nano-biolistics: A method of biolistic transfection of cells and tissues using a gene gun with novel nanometer-sized projectiles. *BMC Biotechnology*, 11(1), 1–6.
- Pearson, H., (2006). What is a gene? *Nature*, 441(7092), 398–402.
- Robl, J., Kasinathan, P., Sullivan, E., Kuroiwa, Y., Tomizuka, K., & Ishida, I., (2003). Artificial chromosome vectors and expression of complex proteins in transgenic animals. *Theriogenology*, 59(1), 107–113.
- Robl, J., Wang, Z., Kasinathan, P., & Kuroiwa, Y., (2007). Transgenic animal production and animal biotechnology. *Theriogenology*, 67(1), 127–133.
- Williamson, J., (2012). *Dragon's Island*. Hachette UK.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

CHAPTER 2

Enzymes of Genetic Engineering

AHMAD ALI¹ and JOHRA KHAN²

¹Department of Life Sciences, University of Mumbai, Vidyanaagari, Santacruz (E), Mumbai, Maharashtra, India

²Department of Medical Laboratory Sciences, College of Applied Medical Sciences, Majmaah University, Majmaah, Saudi Arabia

ABSTRACT

Since genetic engineering has been used to develop new recombinant species of plants, animals, and microorganisms, different enzymes are discovered based on their properties, sources, and functions to be used in the recombination process. The field of genetic engineering uses several tools and techniques. These tools include a wide array of enzymes viz. endonucleases (ER), ligases, methylases, polymerases, etc. In this chapter, we tried to summarize some of these enzymes, their types, and their function used in genetic engineering and molecular biology to help in better understanding and utilization of these enzymes.

2.1 INTRODUCTION

The term “enzyme,” was coined by Wilhelm Kuhne in 1879 during his study of alcohol production using yeast. Enzymes, also known as biocatalyst helps in increasing the speed of all biochemical reactions in all living

organisms (Alotaibi & Khan, 2016; Poulsen et al., 2007). With the development in technology and its application in the production of different products like; antibiotic modification, developing washing powder, sweetening agents, and in the diagnosis of diseases (Boado et al., 2008; Sharma et al., 2018). Enzymes are required in very little quantity to speed up biochemical reactions without getting used up in reaction (Laskin et al., 1987; van Beilen & Li, 2002). In a chemical reaction, the enzyme helps to convert the substrate to the product, as shown in the reaction given in Figure 2.1.

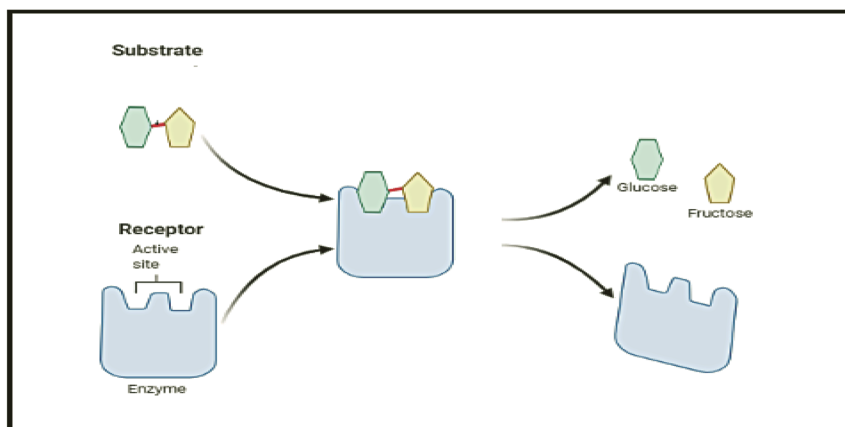


FIGURE 2.1 Enzymes help to convert substrate to products.

The commercial use of enzymes started around the early 20th century when the process of extraction, crystallization, and their catalytic activity associated with protein molecules (Galperin, 2008; Hamza, 2017). Till the 1960s, the general belief about enzymes was considered as protein, but in 1980, it was reported that some RNA molecules can also function as catalysts and are known as ribozymes, which also play an important role in gene expression (Archer & Peberdy, 1997). During this decade, the isolation of antibodies which can function as enzymes in performing catalysis also took place known as abzymes. Abzymes are found substantial as industrial and therapeutics catalysts (Nevinsky & Buneva, 2005). These enzymes are always required in small quantities to increase the rate of reaction without involving in the reaction (Avalle et al., 2000; Janda et al., 1991).

2.2 RESTRICTION ENZYMES: EXO- AND ENDONUCLEASES (ER)

2.2.1 TYPE OF RESTRICTION ENZYMES

A restriction enzyme is also known as “biological scissor” comes under the nuclease enzyme category, which can cleave DNA at specific or random sites as per the requirement and as its identification site (Pingoud et al., 2005; Pingoud & Jeltsch, 2001). HindII was the first restriction enzyme isolated from *Haemophilus influenzae* (Nei & Tajima, 1981; Pingoud & Jeltsch, 2001), followed by the discovery of many restriction endonucleases (ER) by Daniel Nathans, Werner Arber, and Hamilton O. Smith and awarded the Nobile prize of Physiology (Arber & Nathans, 1992). The use of restriction enzymes plays an important role in genetic engineering. Some of the restriction enzymes can identify DNA and RNA at specific sites and cut down phosphodiester bonds. Some of these are; EcoRI, HindIII, and BamHI (Graham et al., 2015; Kim et al., 2016).

EcoRI was the first restriction enzyme isolated in 1968 from *E. coli* and in 1970 second restriction enzyme was HindII isolated from *Haemophilus influenzae* (Beadell & Fleischer, 2005; Lin & O’Callaghan, 2018). Most of the restriction enzymes only identify palindromic sequences in nature. These forms are always mirrored images of each other and always repeated at many places in the genome, like CTAATC, or CCATGG as complementary to GGTACC (Kanaras et al., 2007; Mohammadi et al., 2015). The place where a restriction enzyme cuts a palindromic sequence is known as a restriction site. The cut that these enzymes produce can be either blunt ends as produced by EcoRV, and SmaI or sticky ends like EcoRI, BamHI, and HindIII (Figure 2.2) (Rezaei-Matehkolaei et al., 2012; Tóth et al., 2014). Various restriction enzymes can identify the same restriction sequence but produce cut at different places, such cutting methods are known as neoschizomers like ApaI, and Bsp1201 (Terletskiy et al., 2014). If the identification site and restriction site of two restriction enzymes are the same the method is known as isoschizomers like SmaI and XmaI (Gupta et al., 2012; Ruiz-García et al., 2010; Ryu & Rowsell, 2008).

The classification and nomenclature of the restriction enzyme are done by following some basic rules. The name of restriction endonuclease consists of letters and numbers (Cornish-Bowden, 2014; Makarova et al., 2018; Nørskov-Lauritsen et al., 2019). Some names like Hind II and Hind III consist of three parts as shown in Figure 2.3.

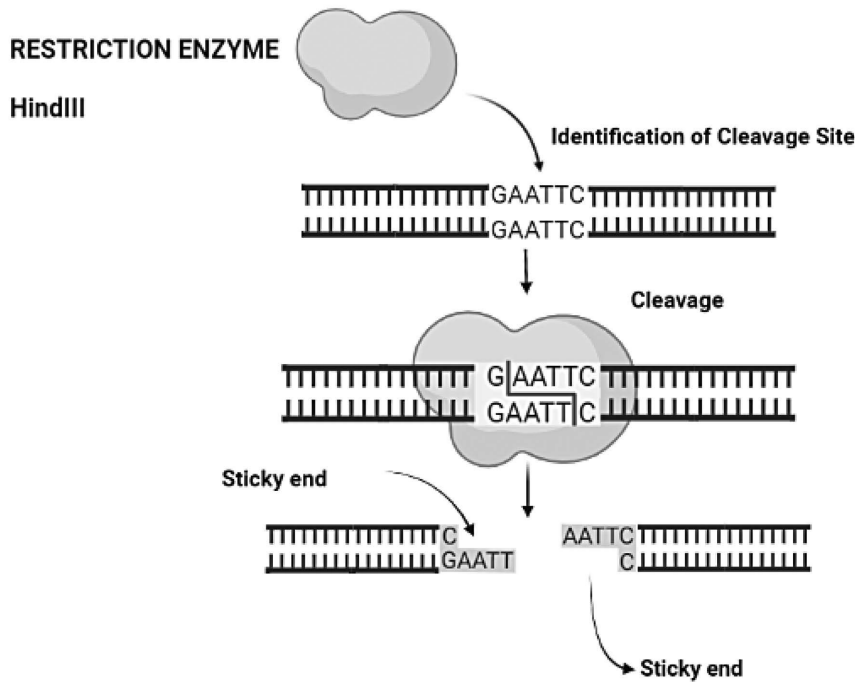


FIGURE 2.2 Restriction enzyme HindIII and its function to produce sticky ends.

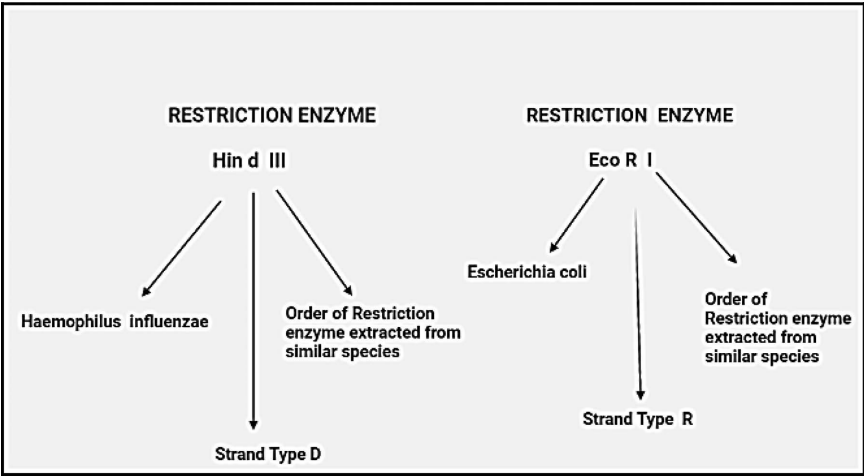


FIGURE 2.3 Nomenclature method of restriction enzyme.

2.2.2 CLASSIFICATION OF RESTRICTION ENZYMES

Restriction enzymes are classified based on their cofactor needs, the composition of subunits, target sequence nature, and place of restriction site of target sequence (Garcia-Bonete et al., 2019; Wilson & Walker, 2010). They are classified into four types as follows:

1. **Type 1:** It includes restriction enzymes that cut the palindrome sequence away from the target sequence in an asymmetric fashion and requires ATP and a cofactor S-adenosyl L-methionine (AdoMet) to work properly. This type can function both as restriction digestion, and methylation (Zumbach et al., 2006).
2. **Type 2:** It includes an enzyme type that cuts near the palindrome sequence. It needs magnesium to function as a cofactor and functions as an independent restriction digestion enzyme (White et al., 2013).
3. **Type 3:** This type includes a restriction enzyme that can recognize 2 distinct non-palindromic forms and non-mirror image sequences and can cut them in size of 20–30 base pairs far from the target site. They consist of more than one subunit of cofactors and need ATP and S-adenosyl-L methionine for DNA Methylation production (Treacy et al., 2011).
4. **Type 4:** This type of restriction enzyme can recognize modified DNA caused due to methylation, glucosyl-hydroxymethylation, and hydroxymethylated DNA. This group includes *E. coli* Mrr and McrBC systems (Artursson et al., 2005; Tropp, 2012).

From the discovery of restriction enzymes, more than 4,000 various restriction enzymes are recorded which can recognize more than 360 restriction sites from different bacterial species (Loenen et al., 2014). Researcher in the field of genetic engineering rely on restriction enzymes for creating new or recombinant DNA (rDNA) in the process of cloning, epigenetics, diagnostics for genetic diseases, creating GMO, paternity testing, forensics, and different field of biotechnology (Bukowski et al., 2015; Treusch et al., 2009). Smith was the first scientist who produced the first synthetic bacterial cell (Gibson et al., 2010). For the synthesis of 1,00,000 base pairs of *Mycoplasma mycoides* Craig Venter used a restriction enzyme to cut 1,080 different base pairs and joined them again. This new genome was then transferred to *M. capricolum* to study the capability

of bacteria to synthesize copies of the *M. mycoide* genome (Choi et al., 2011; Glass, 2012; Meroueh et al., 2006). During the study of restriction enzyme, it was found that the gene responsible for coding of restriction enzyme are found near or next to the methyltransferase coding gene (Gibson et al., 2010; Ruiz-Avila et al., 2013). The restriction enzyme of different bacterial species can have variations in the order and orientation of these genes. In some species, they are found as R and M converge and in some cases as diverge. In species with R and M diverge gene has their codons 50–100 bp distant to gene and in converge condition it has 3–50 bp separated to codons (Paterson et al., 2010; Silva & Antunes, 2017).

2.3 DNA POLYMERASES

In the process of replication of DNA, an enzyme can join the small units like dNTPs (deoxyribonucleotides triphosphates) to create a new continuous strand (Beard & Wilson, 2006). DNA polymerase enzyme starts its function by attaching nucleotides to the 3'OH end of a primer that already contains a 5' phosphate group. DNA polymerase used in genetic engineering and molecular biology is either isolated from bacteria or bacteriophages (virus infecting bacteria) (Sengupta et al., 2014). There are 3 different types of polymerase enzyme found in bacterial cells; Polymerase I, Polymerase II, and Polymerase III. These three enzymes start the processes of isolation from the 5' → 3' direction in the presence of dNTPs and primer during the process of replication. Kornberg and Geftter discovered polymerase III enzyme in 1970 from a complex prokaryotic replication (Barakat et al., 2012; Roether et al., 2019). Polymerase enzymes not only perform forward but are also able to reverse exonuclease activity, which is also a proofreading activity to reduce the chances of errors during replication. DNA polymerase I perform exonuclease activity from 5' → 3' direction to be used for nick-translation which is a tagging method where some nucleotides are replaced by labeled analogs during DNA sequencing (Lange et al., 2011; Moser et al., 2012; Wang et al., 2007).

Polynuclease enzyme is also used to remove protruding DNA at 3' end and sometimes to cover up cohesive ends before adding it to molecular linkers. DNA polymerase I exonuclease activity is not considered suitable for its application in polymerization or to fill cohesive ends before adding them to linkers or before copying of ssDNA by dideoxy method (Batic

et al., 2015; Church et al., 2013). The proteolytic digestion DNA Polymerase I of *E. coli* produces a large DNA Poly IK (Polymerase I Klenow) fragment with 5' → 3' and a short exonuclease fragment for proofreading activities. Klenow and Henningsen improved the quality of these fragments by removing contamination in the presence of residual enzyme treated with proteolytical preparation (Nishida et al., 2009; Shinbrot et al., 2014; Zahurancik & Suo, 2020). T4 polymerase from bacteriophage needs primer to show two activities: (i) polymerase activity from 5' → 3' in presence of dNTPs; (ii) and exonuclease activity from 3' → 5' in absence of dNTPs. T4 DNA polymerase is much better in comparison to *E. coli* DNA polymerase as its exonuclease rate is equal to 40 base/min on dsDNA and 4,000 base/min on ssDNA (Adereth et al., 2005; Pellicanò et al., 2021). The polymerization rate of T4 polymerase is approximately 15,000 nucleotide per minute under standard laboratory conditions. A further modification in bacteriophage also modified T4 polymerase to T7 polymerase. Richardson et al. (1987) consider T7 DNA polymerase as an ideal tool for DNA sequencing and genetic engineering. T7 DNA polymerase consists of 2 protein complexes, including a T7 gene 5 products of 84 kDa, and *E. coli* thioredoxin, which is a product of 12 kDa (Fernandez-Leiro et al., 2015). Thioredoxin protein's primary function is to connect T7 gene 5 to a target primer template which activates the polymerization of many nucleotides without causing any dissociation, whereas the T7 gene 5 helps to provide catalytic properties to the complex protein (Qi & Otting, 2019). T7 DNA polymerase has a polymerization rate 70 times faster than reverse transcriptase of AMV because of which it becomes a better option for radioactive probes and amplification of large DNA molecules (Hariharan & Reha-Krantz, 2005; Wang et al., 2019).

2.3.1 THERMOSTABLE POLYMERASES

The discovery of thermostable polymerase enzyme was recorded around the 1970s but its application and importance was understood after the development of PCR (polymerase chain reaction) and the need for a DNA polymerase enzyme that can be stable at high and fluctuated temperature conditions (Chander et al., 2014; Xu et al., 2017). The Bst thermostable DNA polymerase I was isolated in 1968 from *Bacillus stearothermophilus* (Bst) which can easily grow between 39°C to 70°C with an optimum

temperature of 65°C but gets deactivated at 75°C. Bst polymerase enzyme needs a high amount of Mg^{2+} ion concentration for the high rate of polymerization (Blatter et al., 2013; Moser et al., 2012; Terpe, 2013).

Taq Polymerase is another thermostable polymerase enzyme first isolated in 1976 from a hot spring fountain of Yellowstone National Park bacteria known as *Thermophilus aquaticus* or Taq polymerase (Pech et al., 2017; Sauter & Marx, 2006). Taq polymerase has a half-life of 5 min at 100°C and 40 min at 95°C, which makes them most suitable for PCR processes. Similar to Bst polymerase Taq polymerase also require Mg^{2+} ion at 80°C but in low concentration. Taq polymerase also lacks exonuclease activity due to which chances of error are more in PCR using Taq polymerase because of which it is also known as low fidelity enzyme (Gloeckner et al., 2007). But Taq polymerase can add 45,000 new DNA strands before the occurrence of any error. The modern Taq polymerase is the modified Taq polymerase with better accuracy and stability because of which it is more in use in laboratories than high fidelity Taq (Rigoldi et al., 2018; Turner et al., 2007). Table 2.1 shows different types of DNA polymerase and their functions.

2.4 DNA LIGASES

During DNA replication, one of the new copied fragments is in the form of small pieces known as Okazaki fragments, to attach these fragments, the enzyme that works are the ligase enzyme. DNA ligase is used in many processes of genetic engineering with a restriction enzyme to join sticky and blunt ends of fragments digested using a restriction enzyme. The function of DNA ligase follows three steps: (i) it includes the formation of ligase-adenylate as intermediate, which holds a phosphoramidate bond between lysine and AMP molecule; (ii) AMP molecule is transferred to 5'-phosphate end of DNA nick to create DNA adenylate. Research and data analysis by different scientists shows that DNA ligase of bacteria uses NAD^+ as a cofactor, but DNA ligase of eukaryotic organisms, bacteriophages, and viruses use ATP as cofactor, but later studies confirm the use of both NAD^+ and ATP as cofactors even though both are not same inefficiency (Ghosh et al., 2009). The size of a eukaryotic DNA ligase can be as large as 100 kDa, whereas the size of bacteriophage and chlorella virus is recorded as the smallest of all DNA ligases in size. Eukaryotic

TABLE 2.1 Different Types of DNA Polymerases Used in Genetic Engineering

Type of Polymerase Enzyme	Direction of Activity	Function	References
Polymerase I	1. Polymerase in forwarding direction $5' \rightarrow 3'$ 2. Exonuclease in both forward and reverse direction $5' \rightarrow 3'$ $3' \rightarrow 5'$	Used for DNA replication, primer removal, and nick translation	Johnson et al. (2008); Kubista et al. (2006); Watson (2012)
Polymerase II	1. Polymerase in forward direction $5' \rightarrow 3'$ 2. Exonuclease in one direction only $3' \rightarrow 5'$	DNA replication	Dieci et al. (2007); Eid et al. (2009); Muse et al. (2007)
Polymerase III	1. Polymerase in forward direction $5' \rightarrow 3'$ 2. Exonuclease in one direction only $3' \rightarrow 5'$	DNA replication	Borchert et al. (2006); Dieci et al. (2007)
Klenow fragment of Polynuclease I from <i>E. coli</i>	1. Polymerase in forward direction $5' \rightarrow 3'$ 2. Exonuclease in one direction only $3' \rightarrow 5'$	DNA replication specially to fill the gaps and to avoid exonuclease activity in $3'$ direction	Ginsberg & Che (2014); Ricard (2005)
T4 DNA polymerase from Bacteriophage	1. Polymerase in forward direction $5' \rightarrow 3'$ in the presence of dNTPs. 2. Exonuclease in one direction only $3' \rightarrow 5'$ without dNTPs.	DNA replication specially to fill the gaps and to avoid exonuclease activity in $3'$ direction	Adereth et al. (2005); Hariharan & Reha-Krantz (2005); Qi & Otting (2019)

TABLE 2.1 (Continued)

Type of Polymerase Enzyme	Direction of Activity	Function	References
T7 DNA polymerase from modified bacteriophage	1. Polymerase in forward direction 5' → 3' in the presence of dNTPs and also without dNTPs.	For sequencing	Etson et al. (2010); Luo et al. (2007); Tsai & Johnson (2006)
	2. Exonuclease in one direction only 3' → 5' without dNTPs.		
	3. Modified T7 pol Thioredoxin helps to connect polymerase to target primer templet, so dNTPs dissociation can be controlled.		

DNA ligase contains the OB-fold domain with an additional domain that contains zinc fingers and BRCT. T4 DNA ligase is the most used ligase enzyme in genetic engineering due to its properties of ligating both cohesive and blunt ends, and it can also repair nicks of single-stranded duplex DNA, or a hybrid DNA and RNA, or RNA alone (Hernandez et al., 2020; Johnson et al., 2007; Khan et al., 2015).

DNA ligase from *E. coli* works perfectly on cohesive ends of double-stranded DNA (dsDNA) and blunt ends in the presence of polyethylene glycol, and Ficoll, but they are not efficient on hybrids of DNA-RNA or RNA-RNA. Any of the two cofactors NAD⁺ or ATP forms AMP complex on reacting with ligase, here AMP is linked to the ϵ -amino group of lysine connected by phosphoramidate bond to the active site. AMP moiety activates the phosphate group of DNA on 5'-end, it works as a donor, and nucleophilic attack by 3'-hydroxyl group on active phosphorus acts as acceptor. During this process, AMP is released and a phosphodiester bond is formed for which the energy is driven by breaking 2 high-energy phosphate bonds present in the backbone of DNA with ATP. The optimum temperature for T4 ligase functioning is 16°C under *in vitro* conditions (Didenko, 2011; Stano et al., 2005).

By the end of the 19th century, a thermostable DNA ligase was discovered from *Bacillus stearothermophilus*, *Thermus thermophilus*, *Thermus scotoductus*, and *Rhodothermus marinus*. A thermostable ligase can join two fragments in case they are complementary to each other and the primer can attach the target sequence then the only ligase will be able to function properly and join the dNTPs to form a new copy of the target sequence (Le et al., 2013; Lohman et al., 2011).

2.4.1 RNA LIGASES

RNA ligase's main function is to refix the broken RNA in a cell and it catalyzes ATP-dependent phosphodiester bond present between 5' phosphate to 3' OH of DNA and RNA molecules. Similar to DNA ligases, RNA ligases function also takes place in steps which are: (i) reaction of RNA ligases with ATP, AMP intermediate and pyrophosphate; (ii) transfer of AMP moiety to 5' phosphate of ligase adenylate and breaking of RNA to form intermediate RNA-adenylate (AppRNA); and the last step is (iii) nucleophilic based attack on AppRNA by 3'-OH end of RNA. Like T4

DNA ligase, T4 RNA ligase is the best RNA ligase isolated from T4 infected *E. coli*. T4 RNA ligase is used for many functions including ligation of the polynucleotide and single-stranded RNA with an RNA molecule, fast amplification of cDNA ends, labeling of an RNA molecule to 3'-end of RNA for structure analysis, ligation of oligonucleotide adapter to cDNA, primer extension in PCR, and site mapping (Jayaprakash et al., 2011).

T4 RNA ligase 2 is the recently discovered type of T4 ligase found in bacteriophages. T2 ligase 2 is also able to catalyze both inter and intramolecular strand ligation. T4 RNA ligase 2 is more active to join nicks on double-stranded RNA in comparison to joining single-stranded RNA. T4 RNA ligase 2 enzymes can also be used for optimized ligation for cloning of microRNA (Shuman, 2009; Tomkinson et al., 2006).

2.5 SOME OTHER IMPORTANT ENZYMES OF GENETIC ENGINEERING

2.5.1 NUCLEASES

Nuclease is an enzyme that is used for the breakdown of phosphodiester bonds present in the nucleic acid backbone. It is classified into two broad classes; endonuclease, which breaks down internal bonds of nucleic acid, and exonuclease, which function on ends of nucleic acids (Allan et al., 2012; Kiedrowski et al., 2014). DNA-based deoxyribonucleases break nicks present in DNA and ribonuclease breaks RNA. Nuclease is considered very important in cloning and manipulating DNA and RNA. DNase I (deoxyribonuclease I) is a DNA-based endonuclease enzyme that works on single and dsDNA nicks translation, for digestion of protein, creating random dideoxy sequences, and analysis of DNA-protein interaction in DNase footprinting. DNase I is a glycoprotein isolated from the bovine pancreas in the form of an isoenzyme mixture, due to which great care is needed to separate it from the mixture as RNase are also found as contamination. Thermal denaturation is one of the methods for DNase removal from the mixture using 5 mM EGTA at a temperature of 75°C for 5 min (Sato & Takenaka, 2014).

RNase A is a form of pancreatic ribonuclease which can cleave single-stranded RNA at 3' phosphates, and 3'-phospho-oligonucleotides side with Cp and Up ends. RNase is a very useful enzyme for removing RNA contamination during DNA plasmid formation, removing RNA from DNA/

RNA hybrid of mismatch, and for mutation mapping. Other than RNAs A there are many other ribonucleases used for RNA sequencing, some of them are: (i) ribonuclease Phy I isolated from *Physarum polycephalum*; (ii) RNase CL3 isolated from chicken liver; (iii) ribonuclease Phy M which is purified from *Physarum polycephalum*; (iv) RNase T1 specially used for cleaving RNA at G residue and it is isolated from *Aspergillus oryzae*; and (v) RNase T2 is used for cleaving all phosphodiester bond in RNA and like T1 it is also isolated from *Aspergillus oryzae* (Camilloni et al., 2012; Rosenberg, 2008).

2.5.2 METHYLASES

During infection in a bacterial cell by a bacteriophage, bacterial cell modifies their DNA by adding methyl group, and this addition of methyl group is carried out by methylase enzyme. This enzyme helps in catalyzing the transfer of methyl group to ds DNA from S-adenosyl-methionine. Methylase enzymes are used in molecular cloning specially protect DNA from digestion during cDNA library formation, and for inhibiting restriction enzymes of different sources (Boriack-Sjodin & Swinger, 2016; Roberts et al., 2005).

2.5.3 POLYADENYLATE POLYMERASE OR (POLYA) POLYMERASE

PolyA polymerase enzyme of *E. coli* is used for 3' RNA extension for preparation of priming site for cDNA synthesis using oligo-dT or preparation of RNA for poly-dT-based purification. This enzyme can polymerize different nucleotides from 20 to 2,000. It can also catalyze AMP using ATP, also CTP and UTP can be used as substrate (Bandeled & Osherooff, 2007). Poly A polymerase is a very uncommon enzyme as its optimum condition needs a high concentration of cations like NaCl and Mn^{2+} . Poly A polymerase is sensitive to its poly-A products and does not degrade them, and it also uses a diverse amount of ssRNA as primers (Bilska et al., 2020).

Topoisomerase I is considered a ubiquitous protein used in various processes like DNA replication, supercoiling of DNA at unwinding point, and recombination of DNA protein (Durbecq et al., 2004; Pommier, 2006; Salerno et al., 2010). Initially, topoisomerase was considered useful for

circular DNA preparation or study of nucleosome assembly but later it was discovered that topoisomerase is also useful in catalysis of covalent transfer of ssDNA to heterologous DNA (Aubry et al., 2006; Hanke et al., 2003). The property of topoisomerase to function both as nuclease and ligase is used for the development of cloning vectors (Montaner et al., 2005).

2.6 CONCLUSION

A large number of enzymes and their resources are now known to our scientists and used for different processes like cutting, propagation, recombination of DNA or RNA. For the development of standard procedures in genetic engineering, a better understanding of enzymes and creating standard protocol will help in selecting the good source of enzymes and their sources. The information related to enzymes and their functions will also help in creating their commercial products.

KEYWORDS

- ***Bacillus stearothermophilus***
- **deoxyribonucleotides triphosphates**
- **enzymes**
- **genetic engineering**
- **ligases polymerases**
- **restriction endonucleases**

REFERENCES

- Adereth, Y., Champion, K. J., Hsu, T., & Dammai, V., (2005). Site-directed mutagenesis using Pfu DNA polymerase and T4 DNA ligase. *Biotechniques*, 38(6), 864–868.
- Allan, J., Fraser, R. M., Owen-Hughes, T., & Keszenman-Pereyra, D., (2012). Micrococcal nuclease does not substantially bias nucleosome mapping. *Journal of Molecular Biology*, 417(3), 152–164.

- Alotaibi, A., & Khan, J., (2016). Kinetic study and basal media modification of cellulase enzyme production by mutated strain of *Aspergillus fumigates* (AF1) [Research]. *International Journal of Pharma and Bio Sciences*, 7(1), 193–200. <https://www.ijpbs.net/abstract.php?article=NDg5Ng==> (accessed on 12 October 2022).
- Arber, W., & Nathans, D., (1992). In: Hamilton, O. Smith, (ed.), *Physiology or Medicine, 1971–1980*, 469.
- Archer, D. B., & Peberdy, J. F., (1997). The molecular biology of secreted enzyme production by fungi. *Critical Reviews in Biotechnology*, 17(4), 273–306.
- Artursson, V., Finlay, R. D., & Jansson, J. K., (2005). Combined bromodeoxyuridine immunocapture and terminal-restriction fragment length polymorphism analysis highlights differences in the active soil bacterial metagenome due to *Glomus mosseae* inoculation or plant species. *Environmental Microbiology*, 7(12), 1952–1966.
- Aubry, A., Fisher, L. M., Jarlier, V., & Cambau, E., (2006). First functional characterization of a singly expressed bacterial type II topoisomerase: The enzyme from *Mycobacterium tuberculosis*. *Biochemical and Biophysical Research Communications*, 348(1), 158–165.
- Avalle, B., Friboulet, A., & Thomas, D., (2000). Enzymes and abzymes relationships. *Journal of Molecular Catalysis B: Enzymatic*, 10(1–3), 39–45.
- Bandeled, O. J., & Osherooff, N., (2007). Bioflavonoids as poisons of human topoisomerase II α and II β . *Biochemistry*, 46(20), 6097–6108.
- Barakat, K. H., Gajewski, M. M., & Tuszyński, J. A., (2012). DNA polymerase beta (pol β) inhibitors: A comprehensive overview. *Drug Discovery Today*, 17(15, 16), 913–920.
- Beadell, J. S., & Fleischer, R. C., (2005). A restriction enzyme-based assay to distinguish between avian haemosporidian. *Journal of Parasitology*, 91(3), 683–685.
- Beard, W. A., & Wilson, S. H., (2006). Structure and mechanism of DNA polymerase β . *Chemical Reviews*, 106(2), 361–382.
- Bilska, A., Kusio-Kobialka, M., Krawczyk, P. S., Gewartowska, O., Tarkowski, B., Kobylecki, K., Nowis, D., et al., (2020). Immunoglobulin expression and the humoral immune response is regulated by the non-canonical poly (A) polymerase TENT5C. *Nature Communications*, 11(1), 1–17.
- Blatter, N., Bergen, K., Nolte, O., Welte, W., Diederichs, K., Mayer, J., Wieland, M., & Marx, A., (2013). Structure and function of an RNA-reading thermostable DNA polymerase. *Angewandte Chemie International Edition*, 52(45), 11935–11939.
- Boado, R. J., Zhang, Y., Zhang, Y., Xia, C. F., Wang, Y., & Pardridge, W. M., (2008). Genetic engineering of a lysosomal enzyme fusion protein for targeted delivery across the human blood-brain barrier. *Biotechnology and Bioengineering*, 99(2), 475–484.
- Borchert, G. M., Lanier, W., & Davidson, B. L., (2006). RNA polymerase III transcribes human microRNAs. *Nature Structural & Molecular Biology*, 13(12), 1097–1101.
- Boriack-Sjodin, P. A., & Swinger, K. K., (2016). Protein methyltransferases: A distinct, diverse, and dynamic family of enzymes. *Biochemistry*, 55(11), 1557–1569.
- Bratic, A., Kauppila, T. E., Macao, B., Grönke, S., Siibak, T., Stewart, J. B., Baggio, F., et al., (2015). Complementation between polymerase-and exonuclease-deficient mitochondrial DNA polymerase mutants in genomically engineered flies. *Nature Communications*, 6(1), 1–15.
- Bukowski, M., Polakowska, K., Ilczyszyn, W. M., Sitarska, A., Nytko, K., Kosecka, M., Miedzobrodzki, J., et al., (2015). Species determination within *Staphylococcus* genus

- by extended PCR-restriction fragment length polymorphism of *saoC* gene. *FEMS Microbiology Letters*, 362(1), 1–11.
- Camilloni, C., Robustelli, P., Simone, A. D., Cavalli, A., & Vendruscolo, M., (2012). Characterization of the conformational equilibrium between the two major substrates of RNase A using NMR chemical shifts. *Journal of the American Chemical Society*, 134(9), 3968–3971.
- Chander, Y., Koelbl, J., Puckett, J., Moser, M. J., Klingele, A. J., Liles, M. R., Carrias, A., et al., (2014). A novel thermostable polymerase for RNA and DNA loop-mediated isothermal amplification (LAMP). *Frontiers in Microbiology*, 5, 395.
- Choi, W. S., Ha, D., Park, S., & Kim, T., (2011). Synthetic multicellular cell-to-cell communication in inkjet printed bacterial cell systems. *Biomaterials*, 32(10), 2500–2507.
- Church, D. N., Briggs, S. E., Palles, C., Domingo, E., Kearsley, S. J., Grimes, J. M., Gorman, M., et al., (2013). DNA polymerase ϵ and δ exonuclease domain mutations in endometrial cancer. *Human Molecular Genetics*, 22(14), 2820–2828.
- Cornish-Bowden, A., (2014). Current IUBMB recommendations on enzyme nomenclature and kinetics. *Perspectives in Science*, 1(1–6), 74–87.
- Didenko, V. V., (2011). In situ labeling of DNA breaks and apoptosis by T7 DNA polymerase. In: *DNA Damage Detection in Situ, Ex Vivo, and in Vivo* (pp. 37–48). Springer.
- Dieci, G., Fiorino, G., Castelnovo, M., Teichmann, M., & Pagano, A., (2007). The expanding RNA polymerase III transcriptome. *Trends in Genetics*, 23(12), 614–622.
- Durbecq, V., Desmed, C., Paesmans, M., Cardoso, F., Di Leo, A., Mano, M., Rouas, G., et al., (2004). Correlation between topoisomerase-II α gene amplification and protein expression in HER-2 amplified breast cancer. *International Journal of Oncology*, 25(5), 1473–1479.
- Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., et al., (2009). Real-time DNA sequencing from single polymerase molecules. *Science*, 323(5910), 133–138.
- Eton, C. M., Hamdan, S. M., Richardson, C. C., & Van, O. A. M., (2010). Thioredoxin suppresses microscopic hopping of T7 DNA polymerase on duplex DNA. *Proceedings of the National Academy of Sciences*, 107(5), 1900–1905.
- Fernandez-Leiro, R., Conrad, J., Scheres, S. H., & Lamers, M. H., (2015). Cryo-EM structures of the *E. coli* replicative DNA polymerase reveal its dynamic interactions with the DNA sliding clamp, exonuclease and τ . *Elife*, 4, e11134.
- Galperin, M. Y., (2008). The molecular biology database collection: 2008 update. *Nucleic Acids Research*, 36(suppl_1), D2–D4.
- Garcia-Bonete, M. J., Jensen, M., & Katona, G., (2019). A practical guide to developing virtual and augmented reality exercises for teaching structural biology. *Biochemistry and Molecular Biology Education*, 47(1), 16–24.
- Ghosh, S., Marintcheva, B., Takahashi, M., & Richardson, C. C., (2009). C-terminal phenylalanine of bacteriophage T7 single-stranded DNA-binding protein is essential for strand displacement synthesis by T7 DNA polymerase at a nick in DNA. *Journal of Biological Chemistry*, 284(44), 30339–30349.
- Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R. Y., Algire, M. A., Benders, G. A., et al., (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 329(5987), 52–56.
- Ginsberg, S. D., & Che, S., (2014). Methods and compositions for amplification and detection of microRNAs (miRNAs) and noncoding RNAs (ncRNAs) using the signature

- sequence amplification method (SSAM). *Recent Advances in DNA & Gene Sequences (Formerly Recent Patents on DNA & Gene Sequences)*, 8(1), 2–9.
- Glass, J. I., (2012). Synthetic genomics and the construction of a synthetic bacterial cell. *Perspectives in Biology and Medicine*, 55(4), 473–489.
- Gloeckner, C., Sauter, K. B., & Marx, A., (2007). Evolving a thermostable DNA polymerase that amplifies from highly damaged templates. *Angewandte Chemie International Edition*, 46(17), 3115–3117.
- Graham, C. F., Glenn, T. C., McArthur, A. G., Boreham, D. R., Kieran, T., Lance, S., Manzon, R. G., et al., (2015). Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADS eq). *Molecular Ecology Resources*, 15(6), 1304–1315.
- Gupta, R., Xu, S. Y., Sharma, P., & Capalash, N., (2012). Characterization of MspNI (G/GWCC) and MspNII (R/GATCY), novel thermostable type II restriction endonucleases from *Meiothermus* sp., isoschizomers of AvaII and BstYI. *Molecular Biology Reports*, 39(5), 5607–5614.
- Hamza, T. A., (2017). Bacterial protease enzyme: Safe and good alternative for industrial and commercial use. *Int. J. Chem. Biomol. Sci*, 3(1), 1–10.
- Hanke, T., Ramiro, M. J., Trigueros, S., Roca, J., & Larraga, V., (2003). Cloning, functional analysis and post-transcriptional regulation of a type II DNA topoisomerase from *Leishmania infantum*. A new potential target for anti-parasite drugs. *Nucleic Acids Research*, 31(16), 4917–4928.
- Hariharan, C., & Reha-Krantz, L. J., (2005). Using 2-Aminopurine fluorescence to detect bacteriophage T4 DNA polymerase– DNA complexes that are important for primer extension and proofreading reactions. *Biochemistry*, 44(48), 15674–15684.
- Hernandez, A. J., Lee, S. J., Chang, S., Lee, J. A., Loparo, J. J., & Richardson, C. C., (2020). Catalytically inactive T7 DNA polymerase imposes a lethal replication roadblock. *Journal of Biological Chemistry*, 295(28), 9542–9550.
- Janda, K. D., Weinhouse, M. I., Danon, T., Pacelli, K. A., & Schloeder, D. M., (1991). Antibody bait and switch catalysis: A survey of antigens capable of inducing abzymes with acyl-transfer properties. *Journal of the American Chemical Society*, 113(14), 5427–5434.
- Jayaprakash, A. D., Jabado, O., Brown, B. D., & Sachidanandam, R., (2011). Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. *Nucleic Acids Research*, 39(21), e141.
- Johnson, D. E., Takahashi, M., Hamdan, S. M., Lee, S. J., & Richardson, C. C., (2007). Exchange of DNA polymerases at the replication fork of bacteriophage T7. *Proceedings of the National Academy of Sciences*, 104(13), 5312–5317.
- Johnson, S. A., Dubeau, L., & Johnson, D. L., (2008). Enhanced RNA polymerase III-dependent transcription is required for oncogenic transformation. *Journal of Biological Chemistry*, 283(28), 19184–19191.
- Kanaras, A. G., Wang, Z., Hussain, I., Brust, M., Cosstick, R., & Bates, A. D., (2007). Site-specific ligation of DNA-modified gold nanoparticles activated by the restriction enzyme *StyI*. *Small*, 3(1), 67–70.
- Khan, J., Alotaibi, A., & Deka, M., (2015). Effect of colchicine induced mutation on cellulose enzyme production by *Aspergillus fumigatus*. *World J Pharmaceutical Res.*, 4, 461–471.

- Kiedrowski, M. R., Crosby, H. A., Hernandez, F. J., Malone, C. L., McNamara, H. J. O., & Horswill, A. R., (2014). Staphylococcus aureus Nuc2 is a functional, surface-attached extracellular nuclease. *PloS One*, 9(4), e95574.
- Kim, J. E., Huang, R., Chen, H., You, C., & Zhang, Y. H. P., (2016). Facile construction of random gene mutagenesis library for directed evolution without the use of restriction enzyme in *Escherichia coli*. *Biotechnology Journal*, 11(9), 1142–1150.
- Kubista, M., Andrade, J. M., Bengtsson, M., Forootan, A., Jonák, J., Lind, K., Sindelka, R., et al., (2006). The real-time polymerase chain reaction. *Molecular Aspects of Medicine*, 27(2, 3), 95–125.
- Lange, S. S., Takata, K. I., & Wood, R. D., (2011). DNA polymerases and cancer. *Nature Reviews Cancer*, 11(2), 96–110.
- Laskin, A. I., Mosbach, K., Thomas, D., & Wingard, L. B., (1987). *Enzyme Engineering*, 8.
- Le, Y., Chen, H., Zagursky, R., Wu, J. D., & Shao, W., (2013). Thermostable DNA ligase-mediated PCR production of circular plasmid (PCP) and its application in directed evolution via in situ error-prone PCR. *DNA Research*, 20(4), 375–382.
- Lin, D., & O’Callaghan, C. A., (2018). MetClo: Methylase-assisted hierarchical DNA assembly using a single type IIS restriction enzyme. *Nucleic Acids Research*, 46(19), e113.
- Loenen, W. A., Dryden, D. T., Raleigh, E. A., & Wilson, G. G., (2014). Type I restriction enzymes and their relatives. *Nucleic Acids Research*, 42(1), 20–44.
- Lohman, G. J., Tabor, S., & Nichols, N. M., (2011). DNA ligases. *Current Protocols in Molecular Biology*, 94(1), 3.14. 11–13.14. 17.
- Luo, G., Wang, M., Konigsberg, W. H., & Xie, X. S., (2007). Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7 DNA polymerase. *Proceedings of the National Academy of Sciences*, 104(31), 12610–12615.
- Makarova, K. S., Wolf, Y. I., & Koonin, E. V., (2018). Classification and nomenclature of CRISPR-Cas systems: Where from here? *The CRISPR Journal*, 1(5), 325–336.
- Meroueh, S. O., Bencze, K. Z., Heseck, D., Lee, M., Fisher, J. F., Stemmler, T. L., & Mobashery, S., (2006). Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proceedings of the National Academy of Sciences*, 103(12), 4404–4409.
- Mohammadi, R., Abastabar, M., Mirhendi, H., Badali, H., Shadzi, S., Chadeganipour, M., Pourfathi, P., et al., (2015). Use of restriction fragment length polymorphism to rapidly identify dermatophyte species related to dermatophytosis. *Jundishapur Journal of Microbiology*, 8(6).
- Montaner, B., Castillo-Avila, W., Martinell, M., Öllinger, R., Aymami, J., Giralt, E., & Pérez-Tomás, R., (2005). DNA interaction and dual topoisomerase I and II inhibition properties of the anti-tumor drug prodigiosin. *Toxicological Sciences*, 85(2), 870–879.
- Moser, M. J., DiFrancesco, R. A., Gowda, K., Klingele, A. J., Sugar, D. R., Stocki, S., Mead, D. A., & Schoenfeld, T. W., (2012). Thermostable DNA polymerase from a viral metagenome is a potent RT-PCR enzyme. *PloS One*, 7(6), e38371.
- Muse, G. W., Gilchrist, D. A., Nechaev, S., Shah, R., Parker, J. S., Grissom, S. F., Zeitlinger, J., & Adelman, K., (2007). RNA polymerase is poised for activation across the genome. *Nature Genetics*, 39(12), 1507–1511.
- Nei, M., & Tajima, F., (1981). DNA polymorphism detectable by restriction endonucleases. *Genetics*, 97(1), 145–163.

- Nevinsky, G. A., & Buneva, V. N., (2005). Natural catalytic antibodies—Abzymes. *Catalytic Antibodies*, 503–567.
- Nishida, H., Mayanagi, K., Kiyonari, S., Sato, Y., Oyama, T., Ishino, Y., & Morikawa, K., (2009). Structural determinant for switching between the polymerase and exonuclease modes in the PCNA-replicative DNA polymerase complex. *Proceedings of the National Academy of Sciences*, 106(49), 20693–20698.
- Nørskov-Lauritsen, N., Claesson, R., Jensen, A. B., Åberg, C. H., & Haubek, D., (2019). *Aggregatibacter actinomycetemcomitans*: Clinical significance of a pathobiont subjected to ample changes in classification and nomenclature. *Pathogens*, 8(4), 243.
- Paterson, A. H., Freeling, M., Tang, H., & Wang, X., (2010). Insights from the comparison of plant genome sequences. *Annual Review of Plant Biology*, 61, 349–372.
- Pech, A., Achenbach, J., Jahnz, M., Schülzchen, S., Jarosch, F., Bordusa, F., & Klussmann, S., (2017). A thermostable d-polymerase for mirror-image PCR. *Nucleic Acids Research*, 45(7), 3997–4005.
- Pellicanò, G., Al Mamun, M., Jurado-Santiago, D., Villa-Hernández, S., Yin, X., Giannattasio, M., Lanz, M. C., et al., (2021). Checkpoint-mediated DNA polymerase ϵ exonuclease activity curbing counteracts resection-driven fork collapse. *Molecular Cell*.
- Pingoud, A., & Jeltsch, A., (2001). Structure and function of type II restriction endonucleases. *Nucleic Acids Research*, 29(18), 3705–3727.
- Pingoud, A., Fuxreiter, M., Pingoud, V., & Wende, W., (2005). Type II restriction endonucleases: Structure and mechanism. *Cellular and Molecular Life Sciences*, 62(6), 685–707.
- Pommier, Y., (2006). Topoisomerase I inhibitors: Camptothecins and beyond. *Nature Reviews Cancer*, 6(10), 789–802.
- Poulsen, N., Berne, C., Spain, J., & Kroeger, N., (2007). Silica immobilization of an enzyme through genetic engineering of the diatom *Thalassiosira pseudonana*. *Angewandte Chemie International Edition*, 46(11), 1843–1846.
- Qi, R., & Otting, G., (2019). Mutant T4 DNA polymerase for easy cloning and mutagenesis. *PloS One*, 14(1), e0211065.
- Rezaei-Matehkolaei, A., Makimura, K., Shidfar, M., Zaini, F., Eshraghian, M., Jalalizand, N., Nouripour-Sisakht, S., et al., (2012). Use of single-enzyme PCR-restriction digestion barcode targeting the internal transcribed spacers (ITS rDNA) to identify dermatophyte species. *Iranian Journal of Public Health*, 41(3), 82.
- Ricard, M., (2005). Approaches Used to Better Understand the Functions of the Complex. (AVS). *Marine Chemistry*, 97(3–4), 141–197.
- Rigoldi, F., Donini, S., Redaelli, A., Parisini, E., & Gautieri, A., (2018). Engineering of thermostable enzymes for industrial applications. *APL Bioengineering*, 2(1), 011501.
- Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D., (2005). REBASE—Restriction enzymes and DNA methyltransferases. *Nucleic Acids Research*, 33(suppl_1), D230–D232.
- Roether, J., Chu, K. Y., Willenbacher, N., Shen, A. Q., & Bhalla, N., (2019). Real-time monitoring of DNA immobilization and detection of DNA polymerase activity by a microfluidic nanoplasmonic platform. *Biosensors and Bioelectronics*, 142, 111528.
- Rosenberg, H. F., (2008). RNase A ribonucleases and host defense: An evolving story. *Journal of Leukocyte Biology*, 83(5), 1079–1087.

- Ruiz-Avila, L. B., Huecas, S., Artola, M., Vergoñós, A., Ramírez-Aportela, E., Cercenado, E., Barasoain, I., et al., (2013). Synthetic inhibitors of bacterial cell division targeting the GTP-binding site of FtsZ. *ACS Chemical Biology*, 8(9), 2072–2083.
- Ruiz-García, L., Cabezas, J. A., De María, N., & Cervera, M. T., (2010). Isoschizomers and amplified fragment length polymorphism for the detection of specific cytosine methylation changes. In: *Plant Epigenetics* (pp. 63–74). Springer.
- Ryu, J., & Rowsell, E., (2008). Quick identification of type I restriction enzyme isoschizomers using newly developed pTypeI and reference plasmids. *Nucleic Acids Research*, 36(13), e81–e81.
- Salerno, S., Da Settimo, F., Taliani, S., Simorini, F., La Motta, C., Fornaciari, G., & Marini, A. M., (2010). Recent advances in the development of dual topoisomerase I and II inhibitors as anticancer drugs. *Current Medicinal Chemistry*, 17(35), 4270–4290.
- Sato, S., & Takenaka, S., (2014). Highly sensitive nuclease assays based on chemically modified DNA or RNA. *Sensors*, 14(7), 12437–12450.
- Sauter, K. B., & Marx, A., (2006). Evolving thermostable reverse transcriptase activity in a DNA polymerase scaffold. *Angewandte Chemie International Edition*, 45(45), 7633–7635.
- Sengupta, S., Spiering, M. M., Dey, K. K., Duan, W., Patra, D., Butler, P. J., Astumian, R. D., et al., (2014). DNA polymerase as a molecular motor and pump. *ACS Nano*, 8(3), 2410–2418.
- Sharma, B., Dangi, A. K., & Shukla, P., (2018). Contemporary enzyme based technologies for bioremediation: A review. *Journal of Environmental Management*, 210, 10–22.
- Shinbrot, E., Henninger, E. E., Weinhold, N., Covington, K. R., Göksenin, A. Y., Schultz, N., Chao, H., et al., (2014). Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. *Genome Research*, 24(11), 1740–1750.
- Shuman, S., (2009). DNA ligases: Progress and prospects. *Journal of Biological Chemistry*, 284(26), 17365–17369.
- Silva, L., & Antunes, A., (2017). Vomeronasal receptors in vertebrates and the evolution of pheromone detection. *Annual Review of Animal Biosciences*, 5, 353–370.
- Stano, N. M., Jeong, Y. J., Donmez, I., Tummalapalli, P., Levin, M. K., & Patel, S. S., (2005). DNA synthesis provides the driving force to accelerate DNA unwinding by a helicase. *Nature*, 435(7040), 370–373.
- Terletskiy, V., Tyshchenko, V., Myrzakozha, D., Zhanserkenova, O., Ussenbekov, Y., & Adaev, N. L., (2014). Development of the double digest selective label (DDSL) typing technique and its application to *Staphylococcus aureus* epidemiology. *Sains Malaysiana*, 43(12), 1965–1972.
- Terpe, K., (2013). Overview of thermostable DNA polymerases for classical PCR applications: From molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*, 97(24), 10243–10254.
- Tomkinson, A. E., Vijayakumar, S., Pascal, J. M., & Ellenberger, T., (2006). DNA ligases: Structure, reaction mechanism, and function. *Chemical Reviews*, 106(2), 687–699.
- Tóth, E., Huszár, K., Bencsura, P., Kulcsár, P. I., Vodicska, B., Nyeste, A., Welker, Z., et al., (2014). Restriction enzyme body doubles and PCR cloning: On the general use of type IIs restriction enzymes for cloning. *PloS One*, 9(3), e90896.

- Treacy, D. J., Sankaran, S. M., Gordon-Messer, S., Saly, D., Miller, R., Isaac, S. R., & Kosinski-Collins, M. S., (2011). Implementation of a project-based molecular biology laboratory emphasizing protein structure-function relationships in a large introductory biology laboratory course. *CBE—Life Sciences Education*, 10(1), 18–24.
- Treusch, A. H., Vergin, K. L., Finlay, L. A., Donatz, M. G., Burton, R. M., Carlson, C. A., & Giovannoni, S. J., (2009). Seasonality and vertical structure of microbial communities in an ocean gyre. *The ISME Journal*, 3(10), 1148–1163.
- Tropp, B. E., (2012). *Molecular Biology: Genes to Proteins*. Jones & Bartlett Publishers.
- Tsai, Y. C., & Johnson, K. A., (2006). A new paradigm for DNA polymerase specificity. *Biochemistry*, 45(32), 9675–9687.
- Turner, P., Mamo, G., & Karlsson, E. N., (2007). Potential and utilization of thermophiles and thermostable enzymes in biorefining. *Microbial Cell Factories*, 6(1), 1–23.
- Van, B. J. B., & Li, Z., (2002). Enzyme technology: An overview. *Current Opinion in Biotechnology*, 13(4), 338–344.
- Wang, L., Yu, X., Hu, P., Broyde, S., & Zhang, Y., (2007). A water-mediated and substrate-assisted catalytic mechanism for *Sulfolobus solfataricus* DNA polymerase IV. *Journal of the American Chemical Society*, 129(15), 4731–4737.
- Wang, Y., Li, B., Liu, J., & Zhou, H., (2019). T4 DNA polymerase-assisted upgrade of a nicking/polymerization amplification strategy for ultrasensitive electrochemical detection of watermelon mosaic virus. *Analytical and Bioanalytical Chemistry*, 411(13), 2915–2924.
- Watson, J. D., (2012). *The Polymerase Chain Reaction*. Springer Science & Business Media.
- White, H. B., Benore, M. A., Sumter, T. F., Caldwell, B. D., & Bell, E., (2013). What skills should students of undergraduate biochemistry and molecular biology programs have upon graduation? *Biochemistry and Molecular Biology Education*, 41(5), 297–301.
- Wilson, K., & Walker, J., (2010). *Principles and Techniques of Biochemistry and Molecular Biology*. Cambridge University Press.
- Xu, W., Jiang, W., Wang, J., Yu, L., Chen, J., Liu, X., Liu, L., & Zhu, T. F., (2017). Total chemical synthesis of a thermostable enzyme capable of polymerase chain reaction. *Cell Discovery*, 3(1), 1–10.
- Zahurancik, W. J., & Suo, Z., (2020). Kinetic investigation of the polymerase and exonuclease activities of human DNA polymerase ϵ holoenzyme. *Journal of Biological Chemistry*, 295(50), 17251–17264.
- Zumbach, J., Schmitt, S., Reimann, P., & Starkloff, P., (2006). Learning life sciences: Design and development of a virtual molecular biology learning lab. *Journal of Computers in Mathematics and Science Teaching*, 25(3), 281–300.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

CHAPTER 3

Tools Used in Genetic Engineering

ANKITA SHARMA,¹ AHMAD ALI,¹ and JOHRA KHAN^{2,3}

¹Department of Life Sciences, University of Mumbai, Vidyanaagari, Santacruz (E), Mumbai, Maharashtra, India

²Department of Medical Laboratory Sciences, College of Applied Medical Sciences, Majmaah University, Majmaah, Saudi Arabia

³Health and Basic Sciences Research Center, Majmaah University, Majmaah, Saudi Arabia

ABSTRACT

The use of recombination technique was followed by many civilizations for increasing production in agriculture and animal husbandry, but due to lack of specialized tools, the outcome was less and limited. With the development of genetic engineering, many aspects of our lives are changed and it plays an important role in all fields from agriculture to disease diagnosis and treatment. In this chapter, we try to focus on different tools of genetic engineering that made many developments in modern science.

3.1 INTRODUCTION

Genetic engineering is a tool used in modern biology to change an organism's genetic structure using recombinant DNA (rDNA) technology (Flores, 2019). Conventionally, our ancestors used to regulate breeding

by selecting offspring with required traits by manipulating their genomes indirectly (Oliveira et al., 2017). The definition that is recorded is “The process to modification of one or more genes is known as genetic engineering.” It is also known as genetic modification (Baltimore et al., 2015) or genetic manipulation (Akram et al., 2020).

To develop a preferred phenotype in an organism, a selected gene from identified species is generally added to the genome of chosen plant or animal (Zachariah & Pappachen, 2009). The new DNA that is produced by segregating and replication of the genetic material of interest using rDNA technology (Bekana, n.d.; Muntaha et al., 2016) or by creating it artificially (Mehra et al., 2020). For injecting foreign DNA into a host organism, need to have a construct made (Ashok et al., 2015). The first DNA recombinant was developed by Paul Berg developed in 1972 (Berg & Mertz, 2010) by merging DNA of monkey virus SV40 with lambda virus DNA (Yi, 2008). This method is also used to delete genes, also known as “knock out” genes. The new DNA may be added at random or to a particular location in the genome.

Genetically modified (GM) organisms are organisms that have been created by genetic engineering (GMO) (Holst-Jensen, 2009). In 1973, a bacterium was the first GMO developed by Herbert Boyer and Stanley Cohen (Soutar, 2014). When Rudolf Jaenisch injected foreign DNA into a mouse in 1974, he produced the first GM species (Bera, 2009). Genentech (a biotech firm) was the first organization to concentrate on genetic engineering, and it initiated manufacturing human proteins in 1976. Human insulin was GM in the year 1978, and its production using a bacterium were commercialized in 1982 (Islam, 2013).

The use of rDNA technology started with simplest tasks of cloning small DNA pieces and inserting them in a bacterium that can be used for a vast variety of protein production and using its entire genomes, which can be cloned and inserted from one cell to cell using various techniques developed and used in genetic engineering (McHughen & Smyth, 2008). The technique of molecular cloning of extraneous DNA into bacterial plasmid can be circulated in a bacterial host such as *Escherichia coli* (*E. coli*) is stated as recombinant technology (Baeshen et al., 2014). Plasmids are present in many bacterial organisms and can be conjugated from one organism to another within the same species, resulting in recipient cell transformation.

Plasmids play an important part in molecular cloning of foreign DNA from the year 1970s due to their capability to proliferate in *E. coli*, a well-known model organism. The discovery of restriction endonucleases (ER) enzyme that allows site-specific breaking of DNA, and considered as DNA-modifying enzymes or biological scissors that enabled microbial alteration possible in human development. An overview of the steps involved in DNA cloning.

The general steps for developing plasmid and λ -phage vectors are as follows:

- The restriction enzyme is used to handle both vector and insert DNA. This results in matching ends on both the insert and vector DNA, allowing them to reanneal (Valdez-Cruz et al., 2010).
- The vector and insert DNA restriction digests are mixed and allowed to reanneal.
- The vector is propagated in a host (bacterium) that has been modified for this function after the DNA ends have been sealed with DNA ligase.
- After dissemination, the bacterial culture or phage lysate is transformed into a rDNA library.

Development of vaccines (Leader et al., 2008) and protein therapies including human insulin (Houdebine, 2009), interferon, and human growth hormone (Carter, 2011) have all benefited from rDNA technology. It's also used in the advancement of gene therapy and the production of clotting factors for the treatment of hemophilia. rDNA technology is also used in agriculture for production of GMOs such as Flavr Savr tomatoes (Lewis, 2014), golden rice rich in proteins (Depositario et al., 2009), Bt-cotton (Sadashivappa & Qaim, 2009) to protect plants from ball worms (Qaim & De Janvry, 2005), and many others.

Various tools used in genetic engineering are Restriction enzymes aid in cutting (Roberts, 2005), polymerases aid in synthesizing (Loenen et al., 2014), and ligases aid in binding (Murtola et al., 2010). The vectors aid in the transmission and integration of the desired gene. These are the critical transporters that bring the preferred gene into the target organism, so they are a vital part of the rDNA technology toolkit. In rDNA technology, the most common vectors are plasmids and bacteriophages. The organism into which the rDNA is inserted is known as the host organism (Roberts et al., 2010). The host is the most powerful tool in rDNA technology, as it uses

enzymes to take in the vector that has been engineered with the desired DNA (Engler et al., 2009).

3.2 ORIGIN OF GENETIC ENGINEERING AS A TOOL

Genetic engineering is also used to express the technology used for to modify DNA sequence(s) in genomes by using different methods. Many scientists have pursued research to explain in what way DNA regulates heredity since its finding as the basic unit of heredity and evolution for the basic dogma of molecular biology that explains process of conversation of RNA from DNA and proteins from RNA (Lanigan et al., 2020). The fundamental view of molecular biology is the blueprint information transfer stored in DNA progressive knowledge from DNA to protein (Crick, 1970).

The word “genetic engineering” was coined in 1932 at the Sixth International Congress of Genetics as “the application of genetic concepts to animal and plant breeding.” Scientists have gained a molecular understanding of inheritance thanks to fundamental developments in biology over the last 12 years. Molecular cloning and DNA sequencing, two technically simple and basic techniques, are useful and accurate methods in and of themselves for learning about the structure and function of genes (Biosciences, 1985).

The advent of rDNA technology occurred as a result of the novel application of well-known methods and procedures for analyzing and altering gene structure and organization in complex genomes. The tools and procedures themselves were not revolutionary, despite their revolutionary effect. Rather, it was the new ways in which they were applied that revolutionized biology (Berg & Mertz, 2010).

Some scientists refer to genetic engineering as a set of procedures for cutting and joining different genomes of same or different species together, especially DNA and introducing the subsequent hybrid DNA into an organism to generate new genetic combinations (Mahfouz et al., 2011). In 1974, a meeting was held at the Asilomar Conference Center in California to discuss issues related to rDNA technology. The conference marked a turning point in scientist social consciousness and sense of public duty. A few of the new technology's pioneers saw its profitable potential and recognized private biotechnology companies. Boyer, who

founded Genentech Inc., was one of the first to do so. The business pioneered the use of bacteria to produce human insulin. For diabetics all over the world, GM human insulin has provided a consistent (Haeusler et al., 2013), expandable, and safe supply. The cloning of the human insulin gene was one of the most significant breakthroughs in rDNA technology (Van Dijk et al., 2015).

3.2.1 APPLICATIONS OF GENETIC ENGINEERING TOOLS IN DIFFERENT ASPECTS OF LIFE

3.2.1.1 PHARMACOGENOMICS

Pharmacogenomics, is a branch of science which helps to study the genetic makeup of an individual and its influences on his or her body's reaction to drugs (Shukla, 2020). Pharmaceutical companies use the basic information of pharmacogenomics to develop medicines based on proteins, enzymes, and RNA molecules linked to particular genes and diseases (Nair, 2019). Pharmacogenomics also includes methods for calculating acceptable medication dosages that are more precise. Doctors may use a patient's genetics to assess how well his or her body can absorb and metabolize a drug (Aneesh et al., 2009).

3.2.1.2 DIAGNOSIS AND TREATMENT OF GENETIC DISEASES

Human gene therapy is characterized as the transfer of engineered genetic material into human cells, usually through viral transduction, to treat a disorder or disease (Dunbar et al., 2018). rDNA technology is a procedure in which a gene of interest or a safe gene is introduced into a vector, which can be plasmid, nanostructured, or viral; the latter is the most commonly used due to its efficiency in invading cells and inserting genetic material (Dunbar et al., 2018). Recently, gene therapy has been shown to be effective in the treatment of SCID. The issue can be solved by inserting the right gene into immune system cells. SCID, or Severe Combined Immunodeficiency, is a heritable immunodeficiency that causes the immune system to malfunction (Wilson et al., 2007).

3.2.1.3 IN MEDICINE

1. **Antibodies:** Manipulation of immunoglobulin genes may result in the production of new antibody molecules. Antibodies were generated by joining the variable regions of mouse antibodies to the constant regions of human antibodies (Harding, 2010). Antibodies with altered constant regions have also been produced (Ching et al., 2018). These GM antibodies provide a one-of-a-kind package of reagents for studying the structure-function relationship within the molecule. They may also be useful in the diagnosis and treatment of human diseases (Zeitlin et al., 2011).
2. **Human Serum Albumin (HSA) (Baeshen et al., 2014):** It is the most abundant protein in the blood and is essential for maintaining plasma oncotic pressure and fluid equilibrium between the body's compartments (Langer et al., 2008). As a result, HSA is commonly used in clinics to treat diseases. However, the scarcity of plasma HSA (pHSA) and the associated safety concerns highlight the value of recombinant HSA (rHSA) as a potential alternative (Li, 2014).
3. **Anti-Hemophilic Factors:** Hemophilia A is currently treated with intravenous infusions of factor VIII (f VIII) concentrates (Lambert et al., 2018), which is very expensive and has the risk of producing inhibitors. Gene therapy, on the other hand, may be able to resolve the drawbacks of factor VIII replacement therapy (Kempton & White, 2009). Despite the fact that hemophilia B gene therapy has shown promise in human clinical trials, hemophilia A gene therapy is still a long way off (Hay, 2000). When compared to factor IX, f VIII is a large protein that is difficult to deliver at therapeutic levels using either viral or non-viral vectors (Ma & Carrizosa, 2006). Numerous methods have been used to engineer the f VIII molecule and resolve the barriers preventing long-term and high-level expression in order to increase f VIII gene delivery (Franchini & Lippi, 2010).
4. **Vaccines:** The expression in microbial cells of genes from pathogens that encode surface antigens capable of inducing neutralizing antibodies in the pathogen's host is the simplest application of modern genetic manipulation methods to vaccine production. This method was successfully used to produce a vaccine against

the hepatitis B virus (HBV), which is now widely used (Verma et al., 2016). Allergic research has benefited from the use of rDNA technology, which has provided sequence information and genetic material for the development of new allergy vaccines (Al-Banaa et al., 2019). One popular technique has been to use the information to create synthetic peptides that mimic specific T-cell or B-cell epitopes (Castaman & Linari, 2016). The use of GM allergens to create hypoallergenic vaccines is an alternative approach that can provide a greater and less selective representation of epitopes (Morfini et al., 2013). In recent years, biotechnologists have created a new term called edible vaccine. Edible vaccines are subunit vaccines in which specific genes are inserted into transgenic plants, which are then induced to produce the encoded protein. Potatoes, bananas, cabbage, corn, soybeans, rice, and legumes are among the foods that have been subjected to this treatment (Mannucci & Franchini, 2013). They are a simple to administer, stock, and accept delivery method for patients of various ages that is also cost-efficient. Edible vaccines hold the promise of dramatically reducing diseases such as measles, hepatitis B, cholera, diarrhea, and other diseases, especially in developing countries (Lal et al., 2007). However, in order to make edible vaccines more effective and applicable, a number of technological and regulatory hurdles must be met in the direction of this new vaccine technology (Sahoo et al., 2020).

3.2.1.4 RDNA TECHNOLOGY AND AGRICULTURE

rDNA technologies, which is based on DNA molecular markers, transgenic technology, and gene expression have all been commonly used in agricultural production and have revealed great promise in increasing agricultural harvests and quality (Varshney, 2009), reducing damages caused by various biotic and abiotic stress (Nadeem, 2018), endorsing germplasm resource use, improving breeding performance (Jamil, 2011), and strengthening regulatory oversight. DNA markers have various applications:

- Cultivar identification and seed transparency;
- Assessment of germplasm resource;
- Heterosis forecast.

3.3 GM CROPS

Genetically modified organisms or GMOs denotes the transfer of genetic material between different or same species using various laboratory technique such as gene cloning, splicing DNA pieces together, and introducing genes into different organisms (Varshney, 2009). rDNA technology denotes both of these practices together. Some of the common terms used for these organisms are GMO (Jamil, 2011), genetically engineered (GE) organisms (Nadeem, 2018), bioengineered or transgenic plants or animals are alternative names for GM plants or foods derived from them.

GM crops were first introduced in the United States around 1900 due to their properties like insect resistance or herbicide tolerance, and they were bred in large amount which resulted in existing GM crops grown in the United States (Yamaguchi & Blumwald, 2005). The three most commonly planted GM crops are corn, soybeans, and cotton. An antibiotic-resistant tobacco plant was the first used GMO crop to generate F1 generation of GMO crops. Around the 1990s, China became the first country to introduction of virus-resistant tobacco as commercial transgenic crop. The Food and Drug Administration (FDA) approved the transgenic ‘Flavor Saver tomato’ for sale in the United States in 1994 (Khush, 2013). The tomato was able to postpone ripening after being picked as a result of the update. Few transgenic crops were approved for commercialization in 1995. Several GM crops are currently used as food sources (Tester & Langridge, 2010). There are currently no GM animals that have been licensed for use as food. Genetic engineering has resulted in papaya that is immune to the ring spot virus, increasing productivity. Sugar from GM sugar beets is highly processed and contains no DNA or protein—just its sucrose, just like sugar from non-GM sugar beets. Technologies for genetically engineering crops hold a lot of promise for addressing some of the century’s most pressing issues. They, like all emerging technology, carry threats, both known and unknown. Human and environmental protection, labeling, and consumer preference, intellectual property rights, ethics, food security, poverty reduction, and environmental conservation are all common topics of debate and public concern regarding GM foods and crops (Aneja et al., 2012).

Due to the impossibility of growing traditional cotton in areas with high pest pressure, GM cotton was first commercially planted in Mexico and five other countries in 1996. Since then, 15 countries (Argentina,

Australia, Burma, Brazil, Burkina Faso, China, Colombia, Costa Rica, the United States, India, Mexico, Paraguay, South Africa, Sudan, and Pakistan) have commercialized GM cotton (Blum, 2011).

Bacillus thuringiensis (Bt) is a special bacterium in that it shares a common ancestor with a variety of chemical compounds that are commercially used to regulate insects that are harmful to agriculture and public health (Blaise et al., 2014). Some bacteria, such as *B. popilliae* and *B. sphaericus*, are used as microbial insecticides, but their insecticidal range is much narrower than Bt's (Naik et al., 2005). Importantly, Bt is a human-safe biopesticide and the most commonly used environmentally friendly biopesticide on the planet. Insecticidal Bt genes have also been introduced into many major crops, making them insect resistant and serving as a blueprint for agricultural genetic engineering (Kulvir et al., 2007).

3.3.1 APPLICATION IN INDUSTRY

Reports of effective new biotechnology applications in the food industry give credence to the forecast that biotechnology will be the next major source of competitive advantage at the corporate and international levels (Barrie et al., 2019), and that it will be the most significant technological factor in consolidation strategies. New applications of biotechnology in each of the following food-related areas: enzymes, including the processing of cheese; fermentation, including brewing and winemaking; agricultural raw materials (e.g., crop plants, meat, poultry, fish) with improved functionality; and plant cell bioreactors for food ingredient production. The enzyme alpha-amylase is used in the first step of making high-fructose corn syrup (HFCS), a common nutritive sweetener made from corn-starch (Popa et al., 2019).

Microbial products generated during the exponential phase of growth whose synthesis is an essential part of the normal growth process are known as primary metabolites. They contain anabolic metabolism's intermediates and end products, which the cell uses as building blocks for important macromolecules (e.g., amino acids, nucleotides) or converts to coenzymes (e.g., vitamins). Amino acids, nucleotides, antioxidants, solvents, and organic acids are the most common primary metabolites in industry (Stewart & McLean, 2005).

To grow strains that overproduce primary metabolites, recent methods have used modern genetic engineering techniques. Using *in vitro* rDNA

techniques, this justification for strain construction aims to assemble the required characteristics (Thomas, 2008).

3.4 TOOLS USED IN GENETIC ENGINEERING

3.4.1 ENZYMES INVOLVED IN GENETIC ENGINEERING

Nucleic acids can be modified in a variety of ways to give them unique characteristics and properties. Propagation, ligation, digestion, and the addition of modifying groups such as phosphate or methyl groups are examples of such manipulations (Yuan et al., 2012). Polymerases, ligases, nucleases, phosphatases, and methylases, respectively, catalyze these changes (Metsalu, 2015).

3.4.1.1 RESTRICTION ENZYMES

Restriction enzymes (originally isolated from bacteria) cut double-stranded DNA (dsDNA) into four, five, or six-nucleotide fragments at specific nucleotide sequences (Hummel et al., 2012). A restriction enzyme can cleave DNA wherever it recognizes the enzyme's specific recognition sequence (or restriction site). The end result is a sequence of DNA "restriction fragments." Individual variations in the lengths of these fragments suggest differences in the presence or absence of restriction sites, and studies to investigate length differences are known as restriction fragment length polymorphisms (RFLPs) (Cradick et al., 2010). Restriction enzymes, also known as restriction ER. The majority of restriction enzymes recognize four to eight base pair sequences and hydrolyze one phosphodiester bond on each strand. The double rotational symmetry of many of these cleavage or restriction sites is a distinguishing feature. Cleavage sites are typically symmetrically positioned, or palindromic. Restriction enzymes may produce sticky-end fragments or blunt ends (Liang et al., 2017).

3.4.1.1.1 Nomenclature of Restriction Enzymes

In 1973, Smith and Nathans proposed restriction endonuclease naming guidelines. The enzyme names must begin with an italicized three-letter acronym, according to these guidelines. The first letter represents the

first letter of the bacterial genus from which the enzyme was isolated, while the following two letters represent the bacterial species. To indicate the serotype or strain, these may be accompanied by additional letters or numbers. This is accompanied by a space and a Roman numeral, which indicates the identification's chronological order. Hind III, for example, was the third of four enzymes found in *Haemophilus influenzae* serotype d (Steer et al., 2009).

3.4.1.1.2 *Types of Restriction Enzymes*

There are two different kinds of restriction enzymes:

1. **Exonucleases:** These are a diverse group of enzymes that have been structurally and biochemically characterized (Anastasakis et al., 2016). Exonucleases are enzymes that cleave DNA sequences from the 5' or 3' end of a polynucleotide chain one at a time. *Examples:* Exonuclease I, Exonuclease II, etc. (Shannon et al., 2020).
2. **Endonucleases (ER):** These are enzymes that break a polynucleotide chain's phosphodiester connection. Some, such as deoxyribonuclease I (DNase I), cut DNA in a nonspecific (non-sequence) manner, while others, known as restriction ER or restriction enzymes, cleave only at very unique nucleotide sequences. Exonucleases cleave the ends of recognition sequences rather than the middle (endo)section, while ER cleave the middle portion (Chandler et al., 2013). Mg^{2+} or another divalent cation is needed for the action of restriction ER and many other enzymes that act on phosphate-containing substrates. Restriction ER (restriction enzymes) are classified into three types based on their mechanism of action: Type I, Type II, and Type III (Stoddard, 2005):
 - i. **Type I:** This type DNA restriction/modification (RM) enzymes are bacterial molecular machines that can be found in the majority of bacteria. They maintain sequence-specific methylation (modification) of host DNA while controlling (restricting) the influx of foreign DNA into the bacterium through horizontal gene transfer (HGT) (Pingoud, 2005). Type I restriction-modification (R–M) systems are multifunctional

complexes composed of three different subunits, HsdS (S), HsdM (M) and HsdR (R), which are present in a stoichiometry of 1:2:2. The Type I R–M enzymes require Mg^{2+} , S-adenosyl-methionine, and ATP hydrolysis for the endonuclease activity (Takeuchi et al., 2011). It recognizes specific DNA sequences but cuts at apparently random locations up to 1,000 base pairs away from the recognition site. *Examples:* EcoK1, EcoB1.

- ii. **Type II:** These type enzymes cut DNA at specific locations within or near to their recognition sequences. They are the only class used in the laboratory for routine DNA analysis and gene cloning because they generate discrete restriction fragments and distinct gel banding patterns (Pingoud, 2005). Type II enzymes are a group of unrelated proteins of several different kinds, rather than forming a single family of similar proteins (Schwarz et al., 2013). There have been over 3,000 type II restriction ER discovered. They recognize short, palindromic sequences of 4–8 bp and cleave the DNA inside or near to the recognition sequence in the presence of Mg^{2+} . Type II enzymes often vary so drastically in amino acid sequence from one another, and indeed from any other known protein, that they represent a class of rapidly evolving proteins that are often associated with host-parasite interactions. Type II enzymes cut DNA at specific locations within or near to their recognition sequences (Pingoud et al., 2014). They contain distinct gel banding patterns and isolated restriction fragments. *Examples:* EcoRI, HindIII.
- iii. **Type III:** These type enzymes are also broad restriction-modification enzymes that combine restriction and modification. They seldom offer complete digests because they cleave outside of their recognition sequences, which involve two such sequences in opposite orientations within the same DNA molecule. Type III restriction enzymes can distinguish between two inversely directed non-palindromic sequences. After the recognition site, they cut DNA around 20–30 base pairs (Rao et al., 2014). *Examples:* EcoP1I, Eco15I.

3.4.1.1.3 Staggered Cut/Sticky Ends by Restriction Endonucleases (ER)

After digestion with some restriction enzymes, the ends of the DNA are left with one strand overhanging the other, forming a short (4 nt) single-stranded fragment. This overhang is known as “Sticky ends” because it can easily reattach to other similar ends (Schwarz et al., 2013). R1 restriction endonuclease cleaves duplex DNA by making two single-strand staggered cleavages, producing 5'-phosphoryl and 3'-hydroxyl termini, at a particular series, possibly 6 nucleotide pairs in length (Crampton et al., 2007). Each break produces single-strand ends with similar and complementary sequences of 4 or 6 nucleotides in length. As a result, the cleavage site has a two-fold symmetry axis perpendicular to the helix axis. *E. coli* ligase can join the ends of full-length linear SV40 DNA formed by R1 endonuclease cleavage to regenerate duplex, completely infectious, covalently closed circular molecules. Ends generated by the R1 endonuclease are all identical and complementary. As a result, the sequential action of R1 endonuclease and DNA ligase will “recombine” any two DNA molecules with R1 sites at their restriction sites, resulting in hybrid DNA molecules (Roberts et al., 2005).

3.4.1.1.4 Blunt Ends by Restriction Endonuclease

Sticky ends aren't produced by all restriction enzymes. Some are “blunt cutters,” which slash straight through a target series with no overhang (Figure 3.1).

3.4.1.2 LIGASES

A DNA ligase may re-join DNA fragments that have been separated by a restriction endonuclease or bind them to a new partner (Alomari, 2018). Energy is needed for the reaction, which is given by adding ATP or NAD to the reaction mixture, depending on the form of ligase used. By joining breaks in the phosphodiester backbone of DNA that occur during replication and recombination, as well as a result of DNA damage and repair, DNA ligases play an important role in maintaining genomic

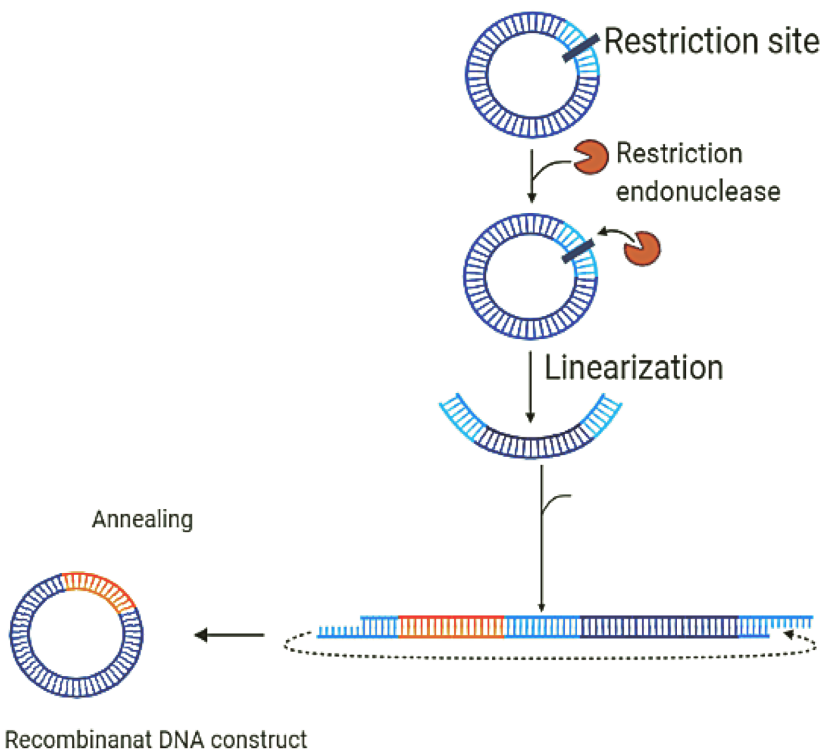


FIGURE 3.1 Use of restriction endonuclease in recombinant DNA construction.

integrity. LIG1, LIG3, and LIG4 are three human genes that code for ATP-dependent DNA ligases. At a single-strand break in dsDNA, DNA ligases catalyze the formation of a phosphodiester bond between a 3'-OH and a 5'-phosphate ring (Malyarchuk et al., 2007).

3.4.1.2.1 Types of Ligases

1. **Non-Thermostable DNA Ligases:** The bacteriophage T4 DNA ligase is the most commonly used ligase in molecular biology. T4 DNA ligase is a 68 kDa monomer that requires the cofactors Mg^{2+} and ATP.

2. **Thermostable DNA Ligases:** At temperatures ranging from 45°C to 80°C, thermostable DNA ligases can ligate duplex molecules and fix single-stranded nicks. They're very precise and well-suited to applications that require high-strength ligations. Thermostable DNA ligases have been isolated from a variety of organisms, including *Thermus thermophilus*. Other examples of ligases – *E. coli* DNA ligase, RNA ligase (Barra & Fachinetti, 2018).

3.4.1.3 KINASES

The Richardson and Hurwitz laboratories discovered polynucleotide kinase (Pnk) in T4 and T2 bacteriophage-infected *E. coli* (Bernstein et al., 2005). The reversible phosphorylation of nucleoside monophosphates, ss (single-stranded) and ds (double-stranded) nucleic acids is catalyzed by Pnk. Pnk, a bacteriophage T4 protein, is commonly used in molecular biology for 5'-phosphorylation of DNA and RNA for labeling and ligation, as well as the removal of 3'-phosphoryl groups from different nucleotide and nucleic acid substrates for labeling purposes. The γ phosphate of ATP is transferred to the 5'-end of nucleic acids by T4 Pnk. There is also a GM variant that lacks phosphatase activity (Tahbaz et al., 2012).

3.4.1.4 PHOSPHATASES

The hydrolase alkaline phosphatase (ALP) extracts the phosphate group from proteins and nucleotides. ALP (EC 3.1.3.1) is a membrane-bound glycoprotein that catalyzes the hydrolysis of phosphate monoesters at basic pH values (McComb et al., 2013).

3.4.1.5 POLYMERASES

In living cells, DNA polymerase is a ubiquitous enzyme that synthesizes complementary DNA (cDNA) strands based on template DNA (Lin et al., 1999). DNA polymerases are commonly used for DNA manipulation *in vitro*, including DNA cloning, sequencing, marking, mutagenesis, and other uses, in addition to their fundamental function in preserving genome integrity during replication and repair. DNA polymerases have the

same basic ability to synthesize a deoxyribonucleotide chain (Al-Hakim et al., 2010). However, the enzymes' more complex characteristics, such as processivity, fidelity (synthesis accuracy), and substrate nucleotide selectivity, vary. In *E. coli*, five DNA polymerases have been identified, eight in *Saccharomyces cerevisiae*, and at least 15 in humans. Although polymerases in bacteria, yeast, and mammalian cells have received a lot of attention, little is known about their plant counterparts. The plant model organism *Arabidopsis thaliana*, for example, is thought to contain 12 DNA polymerases, the majority of whose roles are unknown. DNA polymerases are a class of polymerases that catalyze the synthesis of polydeoxyribonucleotides from mono-deoxyribonucleoside triphosphates (dNTPs), performing the most basic functions *in vivo*, such as DNA replication, repair, and cell differentiation. In a single organism, various forms of DNA polymerases have been identified, such as three (DNA Pol I, II, and III) in *E. coli* or five (DNA Pol α , β , γ , δ , and ϵ) in higher eukaryotes, each of which is thought to perform a specialized *in vivo* role (s). Among the thermostable DNA polymerases, DNA polymerase I from *Thermus aquaticus* (Taq polymerase) is the most well-known representative enzyme. *T. aquaticus* was isolated from Yellowstone National Park in Montana, USA, and Taq polymerase was discovered (Friedberg et al., 2005).

3.4.1.6 REVERSE TRANSCRIPTASE

In 1970, reverse transcriptase was discovered in a number of retroviruses, including the human immunodeficiency virus (Sadashivappa & Qaim) and the avian myeloblastosis virus (AMV) (Preston et al., 1988). The reverse transcriptase enzyme catalyzes the conversion of RNA template molecules into DNA double helixes, making it a valuable tool in molecular biology research (Roberts et al., 1988). Retroviruses contain an enzyme called reverse transcriptase, which transforms the RNA genome carried by the retrovirus particle into dsDNA. To create RNA: DNA hybrid, reverse transcriptase first transcribes a complementary strand of DNA. The hybrid's RNA strand is then degraded by reverse transcriptase or RNase H. After that, the single-stranded DNA is used as a guide for making double-stranded DNA (cDNA). PCR can be used to amplify the cDNA and create several copies once it has been generated. Reverse transcriptase PCR (RT-PCR) is a combination technique that allows genes to be amplified and cloned as intron-free DNA copies from mRNA (Autexier & Lue, 2006).

3.4.1.7 *TERMINAL NUCLEOTIDYL TRANSFERASE*

Exonucleolytic properties of enzymes involved in extracting mismatched pairs at DNA termini are studied using terminal nucleotidyl transferase, which can synthesize a variety of useful polydeoxynucleotides. A small oligodeoxynucleotide serves as a primer for the transferase, which allows for the untemplated synthesis of a polydeoxynucleotide. The structure and concentration of deoxynucleoside triphosphate substrates given to the enzyme are solely responsible for the product's nature and chain lengths (Verma, 1977).

3.5 **VECTORS FOR GENE CLONING**

The process of isolating and expanding a single fragment of DNA into a host organism in order to produce a large number of identical copies is known as molecular cloning (Burke et al., 1987). Transgenesis is a form of experimentation in which a foreign gene is inserted into an organism's genome, accompanied by germ-line transmission of the gene and examination of the progeny's phenotype in higher organisms. Transgenesis may be used to promote the expression of a desired exogenous gene. Vectors assisting in the process is called cloning vectors (Maniatis et al., 1978).

3.5.1 *CLONING VECTORS*

The following four characteristics would make a perfect cloning vehicle:

- Molecules with a low molecular weight;
- Ability to confer on host cells easily selectable phenotypic traits;
- Single sites for a large number of restriction ER, ideally in genes with a phenotype that can be scored; and
- Ability to replicate inside the host cell, resulting in multiple copies of the rDNA molecule being generated and passed on to daughter cells.

3.5.1.1 *PLASMID VECTORS*

Plasmids are dsDNA molecules that are circular and distinct from the chromosomal DNA of a cell. These extrachromosomal DNAs, which are found naturally in bacteria, yeast, and some higher eukaryotic cells,

coexist with their hosts in a parasitic or symbiotic relationship (Maniatis et al., 1978). Plasmids may be as small as a few 1,000 base pairs or as large as 100 kilobases (kb). Plasmid DNA, like the chromosomal DNA of the host cell, is duplicated before each cell division. At least one copy of the plasmid DNA is segregated to each daughter cell during cell division, ensuring that the plasmid is propagated through successive generations of the host cell (Casadaban & Cohen, 1980).

3.5.1.1.1 *Properties*

- i. Plasmids are spherical DNA-based organisms that are much smaller than chromosomes. Plasmids come in a wide range of sizes. The F-plasmid of *E. coli* is around 1% the size of the *E. coli* chromosome and is fairly normal in this regard. The majority of multicopy plasmids are considerably smaller (ColE plasmids are about 10% the size of the F-plasmid). Large plasmids, up to 10% the size of a chromosome, are sometimes discovered, but they are difficult to deal with and only a few have been properly characterized (Andrup & Andersen, 1999).
- ii. The copy number refers to how many copies of the plasmid are present in each bacterial cell. Most plasmids have one or two copies per chromosome, but some small plasmids, such as the ColE plasmids, may have up to 50 or more. The strength of plasmid-borne characteristics, especially antibiotic resistance, is influenced by the number of copies. The more copies of the plasmid per cell, the more copies of the antibiotic resistance genes there will be, and therefore the higher the degree of antibiotic resistance that results.
- iii. Presence of origin of replication (Ori genes).
- iv. In genetic transformation, selectable markers are required for identifying transformed cells containing a gene of interest within the plant cell. A selectable marker must be included since it allows plasmid-containing transformants to be isolated (Andrup & Andersen, 1999).

3.5.1.1.2 *TYPES OF PLASMIDS*

1. **Conjugative Plasmids:** These are horizontally transmissible extra-chromosomal DNA elements present in a variety of naturally

isolated bacteria. While plasmids can provide beneficial genes to their bacterial hosts, they can also be detrimental to their fitness (Hiraga et al., 1986).

2. **Non-Conjugative Plasmids:** These are the ones that cannot pass to other cells without the help of a conjugative system provided by big, so-called conjugative plasmids. Small plasmids are called non-conjugative plasmids. Their M.W. is typically less than 10 Megadaltons. Furthermore, they are multicopy plasmids, which mean that there are normally 10–20 copies present per chromosome. Non-conjugative plasmids, like all other small DNA molecules, are very appealing for basic research (Hiraga et al., 1986).

3.5.1.1.3 USES OF PLASMIDS AS CLONING VECTORS

1. **pBR322:** The vector pBR322 was named using normal vector nomenclature guidelines. Plasmid is represented by the letter “p.” The letter “BR” indicates which laboratory the vector was created in. Bolivar and Rodriguez, two of the scientists who developed the pBR322 cloning vector in 1977, are commemorated in this part of the vector’s name (Balbás et al., 1986). The plasmid was given the number “322” to differentiate it from other vectors they developed in the lab. The plasmid cloning vehicle pBR322 is a 4,362-bp dsDNA plasmid designed to allow for the easy and rapid preparation of cloned rDNA fragments. It has two antibiotic resistance genes, a replication origin, and a number of restriction sites useful for cloning or sub-cloning restriction fragments. When *E. Coli* lacking the resistance genes are successfully transformed, plasmid cloning vectors confer antibiotic resistance. The two antibiotic resistance genes have distinct cloning sites where DNA fragments can be inserted. This resistance gene is inactivated by opening the plasmid with a RE and inserting a compatible DNA segment with DNA ligase. The transformed cells are grown and plated in the presence of the other antibiotic to achieve selection. The plasmid pBR322 was one of the first EK2 multipurpose cloning vectors (10 years ago) to be designed and constructed for efficient cloning and selection of rDNA molecules in *E. coli*. Because of its simplicity

and the availability of its nucleotide sequence, this 4,363-bp DNA molecule has been widely used as a cloning vehicle (Wallace et al., 1981).

2. **Bacteriophage DNA Vectors:** Bacteriophage (phage) is viruses that infect bacteria directly. Phage can be found in a wide range of sizes and shapes. They are divided into two groups based on their morphology and genome type (Sulakvelidze et al., 2001) (Figure 3.2).

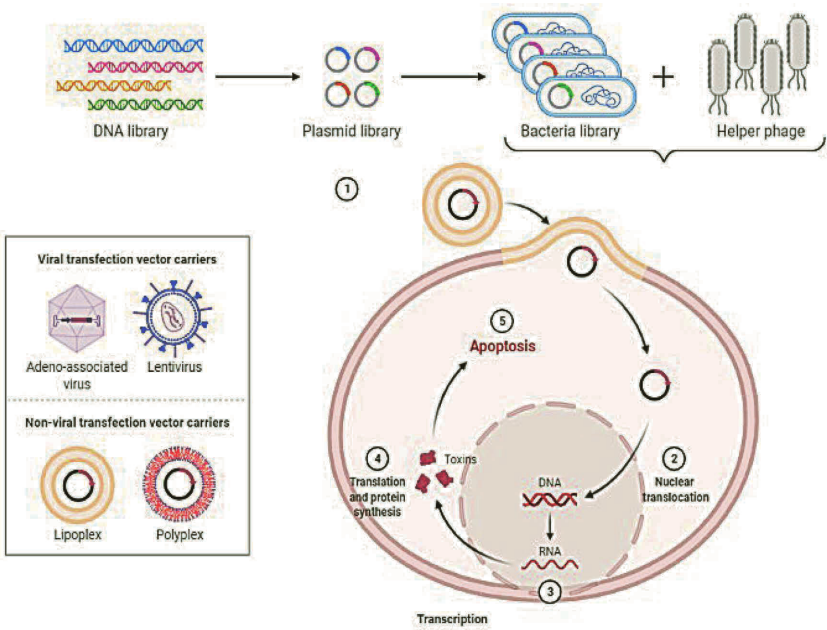


FIGURE 3.2 Different types of phage vectors and their applications.

Phage Lambda as a cloning vector-For a variety of purposes, phage λ is a useful cloning vector. First, phage heads can selectively bundle a chromosome of about 50 kb in length, which can be used to select for molecules with donor DNA inserts, as will be seen. Since the central part of the phage genome isn't required for replication or packaging of DNA molecules in *E. coli*, it can be cut out and discarded using restriction enzymes (d'Herelle & Smith, 1926). The two "arms" are ligated to donor DNA that has been restriction digested. The chimeric molecules may either be

transformed directly into *E. coli* or packaged into phage heads *in vitro*. DNA and phage-head components are mixed together in the *in vitro* environment, and infective λ phages form spontaneously. Recombinant molecules with 10- to 15-kb inserts can be most effectively packed into phage heads, regardless of packaging process, since this size of insert replaces the deleted central part of the phage genome and brings the total molecule size to 50 kb (Monk et al., 2010). As a result, the presence of a phage plaque on the bacterial lawn indicates the presence of a recombinant phage with an insert. A phage vector's second advantage is that recombinant molecules are automatically packed into infective phage particles, which can be processed and treated easily in the lab. Bacteriophage lambda has been used widely as an expression vector and as a cloning vector for over 25 years. Lambda as a cloning vector has the benefits of reliability in packaging and contamination, as well as the ease of plaque screening. A variety of ingenious modifications assist in overcoming the drawbacks associated with its growth mode and scale. Some lambda vectors have been engineered to transform easily into plasmids or phagemids, and a number of promoters and fusions can be used to drive the expression of foreign genes. Antibodies or ligands are used to screen lambda libraries in order to find new genes. The restriction enzymes EcoRI, HindIII, Bam HI, MboI, and BglII produce DNA fragments of about 20 kilobase pairs (kb), which can be cloned using lambda derivatives. There are two types of lambda vectors – replacement vectors and insertion vectors.

3. **Cosmid Vectors:** Cosmids are cloning vectors that are medium in dimension. These vectors have a cloning power of 35–45 kbp. Collins defined the first cosmid vector in 1978. Cosmid vectors are generated by combining the characteristics of plasmid and bacteriophage vectors. The plasmid provides the replication origin, multiple cloning sites (MCS), and selectable marker, whereas the lambda phage provides only the coherent site or cos site area. The cosmid vector is generated by fusing these together. A lambda phage series of around 200 bp is cloned into the cosmid vector. This is made up of the coefficients cosN, cosB, and cosQ (Cheng et al., 2014).

The following steps are involved in cloning a foreign DNA in a cosmid vector: (i) ligation of the foreign DNA between two cos sites; (ii) concatemeric DNA; (iii) *in vitro* packaging to insert the DNA into the phage head to form the matured phage particle; and (iv) transduction of the cloned DNA into *E. coli*. The cosmids are retained as plasmids after entering the host cell. Cosmids are typically held as high-copy plasmids. By integrating the SV40 root of replication, cosmid vectors can also be used as a shuttle vector. The following are the characteristics of a cosmid vector:

- i. *E. coli*-based cloning vector;
- ii. Made up of a plasmid replication origin, a selectable marker gene, and genetic elements from a bacteriophage (cos site);
- iii. Up to 45 kbp cloning capability;
- iv. It's used to construct a genomic library.

3.6 CLONING TECHNIQUES

3.6.1 HOMOPOLYMER TAILING

Many current applications, including massively parallel sequencing, include the amplification of DNA fragments cloned between user-defined 5'- and 3'-end sequences (MPS). HTML-PCR (homopolymer tail-mediated ligation PCR) is a cost-effective high-throughput and robotic PCR system that needs very little starting template and minimal hands-on effort. HTML-PCR begins by adding controlled-length homopolymer tails to the 3'-termini of a double-stranded genomic template. The annealing-assisted ligation of a hybrid oligonucleotide to the template's recessed 5'-ends is made possible by the homopolymer tails. At its 5'-end, the hybrid oligonucleotide has a user-defined sequence. To make the final product, a PCR reaction is used with this primer and a second primer made up of a longer region complementary to the homopolymer tail and fused to a second 5' user-defined series. The user-defined sequences can be modified to work with a wide range of downstream applications (Braña et al., 2014).

3.6.2 LINKERS AND ADAPTORS

Add synthetic linkers or adaptors to blunt-ended cDNA molecules to introduce restriction enzyme sites to their ends, which can then be ligated to compatible termini in a vector. There are two types of synthetic duplex

molecules used for this purpose: (i) linkers, which are duplex molecules with blunt ends on both ends; and (ii) ready-made adaptors, which form short duplexes with one blunt end and one sticky end. The cloning efficiency of linkers can be up to 5-fold higher than that of adaptors, but adaptors are safer and need less manipulations, as will be seen.

3.6.3 REVERSE TRANSCRIPTION

The retroviral replication cycle is characterized by reverse transcription, or the flow of genetic material from RNA to DNA in the opposite direction. When a viral particle reaches the cytoplasm of a target cell, reverse transcription begins. The viral RNA genome joins the cytoplasm as part of a nucleoprotein complex that is still unknown. Reverse transcription creates a linear DNA duplex in the cytoplasm via a complex sequence of steps. Although this DNA is colinear with its RNA template, it contains long terminal repeats (LTRs), which are not found in viral RNA. Reverse transcriptase is an enzyme present in retroviruses that translates the retrovirus's RNA genome into dsDNA. Gene cloning starts with reverse transcription followed by PCR using the messenger RNA, and thus, recognizing the expressed exons of the eukaryotic gene. Execution RT-PCR on an organism under different growth conditions exposes when a gene of interest is expressed and what environment brings gene expression. To compare these two separate conditions, mRNA is digged out from cells growing in both conditions. RT-PCR is done with the two samples of mRNA using PCR primers that match a particular gene of interest. When the gene is expressed, PCR product will be formed, whereas if the gene is switched off, its mRNA will be missing (Wong, 2006).

3.6.4 GENOMIC DNA LIBRARY

A genomic library is a collection of overlapping genomic DNA segments cloned into a backbone vector that statistically contains all regions of an organism's genome. Following that, the cloned DNA is transformed into a suitable host cell line. Many genetic studies, as well as the isolation and cloning of genes from an organism, begin with the development of a genomic library. Genomic library screening has been useful in identifying genes of interest to the medical and biotechnology industries, as well as genes associated with specific cellular functions. Additionally, genomic

mapping or full genome sequencing includes the development of a representative genomic library of an organism (Wong, 2006).

3.6.5 PCR

Kary Mullis invented PCR in the 1980s, and it was a game-changing system. The ability of DNA polymerase to synthesize new strands of DNA complementary to the template strand is used in PCR. Since DNA polymerase can only add a nucleotide to a 3'-OH group that already exists, it requires a primer to add the first nucleotide. This criterion allows the researcher to define a particular region of the template sequence that he or she needs to amplify. The same sequence will be accumulated in billions of copies at the end of the PCR reaction. Template DNA, primers, nucleotides, and DNA polymerase are all required for each PCR assay. The DNA polymerase enzyme is responsible for joining individual nucleotides to form the PCR product. The nucleotides are the four bases present in DNA: adenine, thymine, cytosine, and guanine (A, T, C, G). These serve as the building blocks for the DNA polymerase to use in creating the final PCR product. The reaction's primers determine the DNA product should be amplified (Namouchi & Mardassi, 2006). Short DNA fragments with a given sequence complementary to the target DNA to be detected and amplified are used as primers. These serve as a foundation for the DNA polymerase to build upon. The above-mentioned components are combined in a test tube or 96-well plate, which is then put in a system that allows for three simple steps of DNA amplification. In essence, the unit is a thermal cycler. The test tubes or plates containing the PCR reaction mixture are placed into a thermal block with holes. The system uses discrete, precise, and pre-programmed steps to raise and lower the temperature of the block. The reaction solution is heated above the melting point of the target DNA's two cDNA strands, allowing the strands to separate in a process known as denaturation. Hybridization or annealing is the process of lowering the temperature to allow the specific primers to bind to the target DNA segments. Annealing between primers and target DNA occurs only when their sequences are complementary (e.g., A binding to G). When the temperature is raised again, the DNA polymerase will expand the primers by adding nucleotides to the growing DNA strand. The number of copied DNA molecules doubles with each repeat of these three steps (Bouzidi et al., 2006).

3.7 CONCLUSION

rDNA technology is a significant advancement in science that has made life much simpler for humans. It has advanced strategies for biomedical applications such as cancer treatment, genetic diseases, diabetes, and a variety of plant disorders, especially viral and fungal resistance, in recent years. For propagating, cutting, or altering nucleic acids, scientists now have access to a vast array of enzymes. A thorough understanding of the function of each enzyme used in a protocol would undoubtedly aid in the selection of the most suitable source of enzymatic activity for a given reason, as well as the definition of the best reaction conditions. Successful cloning experiments require carefully selected enzymatic activities and high-quality reagents.

Different types of plasmids are used for successful gene cloning. Over the years, there have been developments in plasmids. Various different bacterial sources are used for plasmids. The method of copying DNA fragments that can then be used for a variety of purposes, such as developing GM crops or discovering a disease cure. There are two forms of gene cloning: *in vivo*, which involves using restriction enzymes and ligases with vectors and cloning the fragments into host cells, and *ex vivo*, which involves using restriction enzymes and ligases with vectors and cloning the fragments into host cells. The other kind is *in vitro*, which involves creating copies of DNA fragments using the polymerase chain reaction (PCR) process.

KEYWORDS

- **cosmid vectors**
- **endonucleases**
- **hepatitis b virus**
- **pharmacogenomics**
- **plasmid vectors**
- **polymerase chain reaction**
- **recombinant DNA technology**

REFERENCES

- Akram, M., Jabeen, F., Daniyal, M., Zainab, R., Ul Haq, U., Adetunji, C. O., Egbuna, C., et al., (2020). Genetic engineering of novel products of health significance: Recombinant DNA technology. In: *Functional Foods and Nutraceuticals* (pp. 595–611). Springer.
- Al-Banaa, K., Alhillan, A., Hawa, F., Mahmood, R., Zaki, A., El Abdallah, M., ELJack, A., et al., (2019). Eemicizumab use in treatment of acquired hemophilia A: A case report. *The American Journal of Case Reports*, 20, 1046.
- Al-Hakim, A., Escribano-Diaz, C., Landry, M. C., O'Donnell, L., Panier, S., Szilard, R. K., & Durocher, D., (2010). The ubiquitous role of ubiquitin in the DNA damage response. *DNA Repair*, 9(12), 1229–1240.
- Alomari, A., (2018). *Biophysical and Kinetic Analysis of Escherichia coli DNA Ligase Activity and Inhibition*. University of Portsmouth].
- Anastasakis, D., Skeparnias, I., Shaukat, A. N., Grafanaki, K., Kanellou, A., Taraviras, S., Papachristou, D. J., et al., (2016). Mammalian PNLDC1 is a novel poly (A) specific exonuclease with discrete expression during early development. *Nucleic Acids Research*, 44(18), 8908–8920.
- Andrup, L., & Andersen, K., (1999). A comparison of the kinetics of plasmid transfer in the conjugation systems encoded by the F plasmid from *Escherichia coli* and plasmid pCF10 from *Enterococcus faecalis*. *Microbiology*, 145(8), 2001–2009.
- Aneesh, T., Sekhar, S., Jose, A., Chandran, L., & Zachariah, S. M., (2009). Pharmacogenomics: The right drug to the right person. *Journal of Clinical Medicine Research*, 1(4), 191.
- Aneja, B., Yadav, N. R., Chawla, V., & Yadav, R. C., (2012). Sequence-related amplified polymorphism (SRAP) molecular marker system and its applications in crop improvement. *Molecular Breeding*, 30(4), 1635–1648.
- Ashok, N. A., Spoorthi, B. K., Ashcheulova, T., & Gerasimchuk, N., (2015). *Genetic Engineering and Ethical Issues (Doctoral dissertation)*.
- Autexier, C., & Lue, N. F., (2006). The structure and function of telomerase reverse transcriptase. *Annu. Rev. Biochem.*, 75, 493–517.
- Baeshen, N. A., Baeshen, M. N., Sheikh, A., Bora, R. S., Ahmed, M. M. M., Ramadan, H. A., Saini, K. S., & Redwan, E. M., (2014). Cell factories for insulin production. *Microbial Cell Factories*, 13(1), 1–9.
- Balbás, P., Soberón, X., Merino, E., Zurita, M., Lomeli, H., Valle, F., Flores, N., & Bolivar, F., (1986). Plasmid vector pBR322 and its special-purpose derivatives: A review. *Gene*, 50(1–3), 3–40.
- Baltimore, D., Berg, P., Botchan, M., Carroll, D., Charo, R. A., Church, G., Corn, J. E., et al., (2015). A prudent path forward for genomic engineering and germline gene modification. *Science*, 348(6230), 36–38.
- Barra, V., & Fachinetti, D., (2018). The dark side of centromeres: Types, causes and consequences of structural abnormalities implicating centromeric DNA. *Nature Communications*, 9(1), 1–17.
- Barrie, J., Zawdie, G., & João, E., (2019). Assessing the role of triple helix system intermediaries in nurturing an industrial biotechnology innovation network. *Journal of Cleaner Production*, 214, 209–223.

- Bekana, G., (2017). Review on Application of Genetic Engineering to Crop Improvement and Its Perceptions in Ethiopia. *Advances in Life Science and Technolog.* 60, 1–7.
- Bera, R. K., (2009). The story of the Cohen–Boyer patents. *Current Science*, 96(6), 760–763.
- Berg, P., & Mertz, J. E., (2010). Personal reflections on the origins and emergence of recombinant DNA technology. *Genetics*, 184(1), 9–17.
- Bernstein, N. K., Williams, R. S., Rakovszky, M. L., Cui, D., Green, R., Karimi-Busheri, F., Mani, R. S., et al., (2005). The molecular architecture of the mammalian DNA repair enzyme, polynucleotide kinase. *Molecular Cell*, 17(5), 657–670.
- Blaise, D., Venugopalan, M., & Raju, A., (2014). Introduction of Bt cotton hybrids in India: Did it change the agronomy? *Indian Journal of Agronomy*, 59(1), 1–20.
- Blum, A., (2011). Drought resistance—is it really a complex trait? *Functional Plant Biology*, 38(10), 753–757.
- Bouzidi, M. F., Franchel, J., Tao, Q., Stormo, K., Mraz, A., Nicolas, P., & Mouzeyar, S., (2006). A sunflower BAC library suitable for PCR screening and physical mapping of targeted genomic regions. *Theoretical and Applied Genetics*, 113(1), 81–89.
- Braña, A. F., Rodríguez, M., Pahari, P., Rohr, J., García, L. A., & Blanco, G., (2014). Activation and silencing of secondary metabolites in *Streptomyces albus* and *Streptomyces lividans* after transformation with cosmids containing the thienamycin gene cluster from *Streptomyces cattleya*. *Archives of Microbiology*, 196(5), 345–355.
- Burke, D. T., Carle, G. F., & Olson, M. V., (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science*, 236(4803), 806–812.
- Carter, P. J., (2011). Introduction to current and future protein therapeutics: A protein engineering perspective. *Experimental Cell Research*, 317(9), 1261–1269.
- Casadaban, M. J., & Cohen, S. N., (1980). Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *Journal of Molecular Biology*, 138(2), 179–207.
- Castaman, G., & Linari, S., (2016). Human von Willebrand factor/factor VIII concentrates in the management of pediatric patients with von Willebrand disease/hemophilia A. *Therapeutics and Clinical Risk Management*, 12, 1029.
- Chandler, M., De La Cruz, F., Dyda, F., Hickman, A. B., Moncalian, G., & Ton-Hoang, B., (2013). Breaking and joining single-stranded DNA: The HUH endonuclease superfamily. *Nature Reviews Microbiology*, 11(8), 525–538.
- Cheng, J., Pinnell, L., Engel, K., Neufeld, J. D., & Charles, T. C., (2014). Versatile broad-host-range cosmids for construction of high quality metagenomic libraries. *Journal of Microbiological Methods*, 99, 27–34.
- Ching, K. H., Collarini, E. J., Abdiche, Y. N., Bedinger, D., Pedersen, D., Izquierdo, S., Harriman, R., et al., (2018). Chickens with humanized immunoglobulin genes generate antibodies with high affinity and broad epitope coverage to conserved targets. *mAbs*.
- Cradick, T. J., Keck, K., Bradshaw, S., Jamieson, A. C., & McCaffrey, A. P., (2010). Zinc-finger nucleases as a novel therapeutic strategy for targeting hepatitis B virus DNAs. *Molecular Therapy*, 18(5), 947–954.
- Crampton, N., Yokokawa, M., Dryden, D. T., Edwardson, J. M., Rao, D. N., Takeyasu, K., Yoshimura, S. H., & Henderson, R. M., (2007). Fast-scan atomic force microscopy reveals that the type III restriction enzyme EcoP15I is capable of DNA translocation and looping. *Proceedings of the National Academy of Sciences*, 104(31), 12755–12760.

- D'Herelle, F., & Smith, G. H., (1926). *The Bacteriophage and Its Behavior*. Am Assoc Immunol.
- Depositario, D. P. T., Nayga, J., Rodolfo, M., Wu, X., & Laude, T. P., (2009). Effects of information on consumers' willingness to pay for golden rice. *Asian Economic Journal*, 23(4), 457–476.
- Dunbar, C. E., High, K. A., Joung, J. K., Kohn, D. B., Ozawa, K., & Sadelain, M., (2018). Gene therapy comes of age. *Science*, 359(6372).
- Engler, C., Gruetzner, R., Kandzia, R., & Marillonnet, S., (2009). Golden gate shuffling: A one-pot DNA shuffling method based on type II restriction enzymes. *PloS One*, 4(5), e5553.
- Flores, J. L., (2019). *Genetic Engineering in Agriculture*. Library of Parliament= Bibliothèque du Parlement.
- Franchini, M., & Lippi, G., (2010). Von Willebrand factor-containing factor VIII concentrates and inhibitors in haemophilia A. *Thrombosis and Haemostasis*, 104(11), 931–940.
- Friedberg, E. C., Lehmann, A. R., & Fuchs, R. P., (2005). Trading places: How do DNA polymerases switch during translesion DNA synthesis? *Molecular Cell*, 18(5), 499–505.
- Haeusler, R. A., Astiarraga, B., Camastra, S., Accili, D., & Ferrannini, E., (2013). Human insulin resistance is associated with increased plasma levels of 12 α -hydroxylated bile acids. *Diabetes*, 62(12), 4184–4191.
- Hiraga, S., Jaffe, A., Ogura, T., Mori, H., & Takahashi, H., (1986). F plasmid ccd mechanism in *Escherichia coli*. *Journal of Bacteriology*, 166(1), 100–104.
- Holst-Jensen, A., (2009). Testing for genetically modified organisms (GMOs): Past, present and future perspectives. *Biotechnology Advances*, 27(6), 1071–1082.
- Houdebine, L. M., (2009). Production of pharmaceutical proteins by transgenic animals. *Comparative Immunology, Microbiology and Infectious Diseases*, 32(2), 107–121.
- Hummel, A. W., Doyle, E. L., & Bogdanove, A. J., (2012). Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. *New Phytologist*, 195(4), 883–893.
- Islam, M., (2013). Animal models of diabetic neuropathy: Progress since 1960s. *Journal of Diabetes Research*, 2013.
- Kempton, C. L., & White, G. C., (2009). How we treat a hemophilia A patient with a factor VIII inhibitor. *Blood, The Journal of the American Society of Hematology*, 113(1), 11–17.
- Khush, G. S., (2013). Strategies for increasing the yield potential of cereals: Case of rice as an example. *Plant Breeding*, 132(5), 433–436.
- Kulvir, S., Vikas, J., Vikram, S., & Rathore, P., (2007). Performance of Bt cotton hybrids under different geometrical arrangements. *Journal of Cotton Research and Development*, 21(1), 41–44.
- Lal, P., Ramachandran, V., Goyal, R., & Sharma, R., (2007). Edible vaccines: Current status and future. *Indian J. Med. Microbiol.*, 25(2), 93–102.
- Lambert, T., Benson, G., Dolan, G., Hermans, C., Jiménez-Yuste, V., Ljung, R., Morfini, M., et al., (2018). Practical aspects of extended half-life products for the treatment of haemophilia. *Therapeutic Advances in Hematology*, 9(9), 295–308.

- Langer, K., Anhorn, M., Steinhauser, I., Dreis, S., Celebi, D., Schrickel, N., Faust, S., & Vogel, V., (2008). Human serum albumin (HSA) nanoparticles: Reproducibility of preparation process and kinetics of enzymatic degradation. *International Journal of Pharmaceutics*, 347(1, 2), 109–117.
- Leader, B., Baca, Q. J., & Golan, D. E., (2008). Protein therapeutics: A summary and pharmacological classification. *Nature Reviews Drug Discovery*, 7(1), 21–39.
- Lewis, J., (2014). From Flavr Savr tomatoes to stem cell therapy: Young people's understandings of gene technology, 15 years on. *Science & Education*, 23(2), 361–379.
- Liang, X., Potter, J., Kumar, S., Ravinder, N., & Chesnut, J. D., (2017). Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. *Journal of Biotechnology*, 241, 136–146.
- Lin, W., Wu, X., & Wang, Z., (1999). A full-length cDNA of hREV3 is predicted to encode DNA polymerase ζ for damage-induced mutagenesis in humans. *Mutation Research/ DNA Repair*, 433(2), 89–98.
- Loenen, W. A., Dryden, D. T., Raleigh, E. A., Wilson, G. G., & Murray, N. E., (2014). Highlights of the DNA cutters: A short history of the restriction enzymes. *Nucleic Acids Research*, 42(1), 3–19.
- Ma, A. D., & Carrizosa, D., (2006). Acquired factor VIII inhibitors: Pathophysiology and treatment. *Hematology*, 2006(1), 432–437.
- Mahfouz, M. M., Li, L., Shamimuzzaman, M., Wibowo, A., Fang, X., & Zhu, J. K., (2011). De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proceedings of the National Academy of Sciences*, 108(6), 2623–2628.
- Malyarchuk, S., Wright, D., Castore, R., Klepper, E., Weiss, B., Doherty, A. J., & Harrison, L., (2007). Expression of *Mycobacterium tuberculosis* Ku and ligase D in *Escherichia coli* results in RecA and RecB-independent DNA end-joining at regions of microhomology. *DNA Repair*, 6(10), 1413–1424.
- Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K., & Efstratiadis, A., (1978). The isolation of structural genes from libraries of eucaryotic DNA. *Cell*, 15(2), 687–701.
- Mannucci, P. M., & Franchini, M., (2013). Present and future challenges in the treatment of haemophilia: A clinician's perspective. *Blood Transfusion*, 11(Suppl 4), s77.
- McComb, R. B., Bowers, Jr. G. N., & Posen, S., (2013). *Alkaline Phosphatase*. Springer Science & Business Media.
- McHughen, A., & Smyth, S., (2008). US regulatory system for genetically modified [genetically modified organism (GMO), rDNA or transgenic] crop cultivars. *Plant Biotechnology Journal*, 6(1), 2–12.
- Mehra, S., Sharma, P., & Chadha, P., (2020). Chapter 5; Genetically modified organisms: Applications and related ethical issues. *Biotechnology Bioscience*, 65.
- Monk, A., Rees, C., Barrow, P., Hagens, S., & Harper, D., (2010). Bacteriophage applications: Where are we now? *Letters in Applied Microbiology*, 51(4), 363–369.
- Morfini, M., Coppola, A., Franchini, M., & Di Minno, G., (2013). Clinical use of factor VIII and factor IX concentrates. *Blood Transfusion*, 11(Suppl 4), s55.
- Muntaha, S. T., Ahmed, A., & Ahmed, K., (2016). Applications and future prospects of genetic engineering: A new global perspective. *FUUAST Journal of Biology*, 6(2), 201–209.

- Murtola, M., Wenska, M., & Strömberg, R., (2010). PNAzymes that are artificial RNA restriction enzymes. *Journal of the American Chemical Society*, 132(26), 8984–8990.
- Naik, G., Qaim, M., Subramanian, A., & Zilberman, D., (2005). Bt cotton controversy: Some paradoxes explained. *Economic and Political Weekly*, 1514–1517.
- Nair, A. S., (2019). Pharmacogenomics of inhalational anesthetic agents. *Medical Gas Research*, 9(1), 52.
- Namouchi, A., & Mardassi, H., (2006). A genomic library-based amplification approach (GL-PCR) for the mapping of multiple IS6110 insertion sites and strain differentiation of *Mycobacterium tuberculosis*. *Journal of Microbiological Methods*, 67(2), 202–211.
- Oliveira, C., Aguiar, T. Q., & Domingues, L., (2017). Principles of genetic engineering. In: *Current Developments in Biotechnology and Bioengineering* (pp. 81–127). Elsevier.
- Pingoud, A., Wilson, G. G., & Wende, W., (2014). Type II restriction endonucleases—A historical perspective and more. *Nucleic Acids Research*, 42(12), 7489–7527.
- Popa, M. E., Mitelut, A. C., Popa, E. E., & Matei, F., (2019). Creating products and services in food biotechnology. In: *Introduction to Biotech Entrepreneurship: From Idea to Business* (pp. 141–178). Springer.
- Preston, B. D., Poiesz, B. J., & Loeb, L. A., (1988). Fidelity of HIV-1 reverse transcriptase. *Science*, 242(4882), 1168–1171.
- Qaim, M., & De Janvry, A., (2005). Bt cotton and pesticide use in Argentina: Economic and environmental effects. *Environment and Development Economics*, 179–200.
- Rao, D. N., Dryden, D. T., & Bheemanaik, S., (2014). Type III restriction-modification enzymes: A historical perspective. *Nucleic Acids Research*, 42(1), 45–55.
- Roberts, J. D., Bebenek, K., & Kunkel, T. A., (1988). The accuracy of reverse transcriptase from HIV-1. *Science*, 242(4882), 1171–1173.
- Roberts, R. J., (2005). How restriction enzymes became the workhorses of molecular biology. *Proceedings of the National Academy of Sciences*, 102(17), 5905–5908.
- Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D., (2005). REBASE—Restriction enzymes and DNA methyltransferases. *Nucleic Acids Research*, 33(suppl_1), D230–D232.
- Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D., (2010). REBASE—A database for DNA restriction and modification: Enzymes, genes and genomes. *Nucleic Acids Research*, 38(suppl_1), D234–D236.
- Sadashivappa, P., & Qaim, M., (2009). Bt Cotton in India: Development of Benefits and the Role of Government Seed Price Interventions. *AgBioForum*, 12(2) 172–183.
- Sahoo, A., Mandal, A. K., Dwivedi, K., & Kumar, V., (2020). A cross-talk between the immunization and edible vaccine from natural origin: Current challenges and future prospects. *Life Sciences*, 118343.
- Schwarz, F. W., Tóth, J., Van, A. K., Cui, G., Clausing, S., Szczelkun, M. D., & Seidel, R., (2013). The helicase-like domains of type III restriction enzymes trigger long-range diffusion along DNA. *Science*, 340(6130), 353–356.
- Shannon, A., Le, N. T. T., Selisko, B., Eydoux, C., Alvarez, K., Guillemot, J. C., Decroly, E., et al., (2020). Remdesivir and SARS-CoV-2: Structural requirements at both nsp12 RdRp and nsp14 Exonuclease active-sites. *Antiviral Research*, 178, 104793.
- Shukla, R., (2020). Pharmacogenomics: Overview, applications, and recent developments. *Drug Design-Novel Advances in the Omics Field and Applications*.
- Soutar, A., (2014). Herbert Boyer and Stanley Cohen: Recombinant DNA. *Pioneers of Medicine Without a Nobel Prize*, 179.

- Steer, P. A., Kirkpatrick, N. C., O'Rourke, D., & Noormohammadi, A. H., (2009). Classification of fowl adenovirus serotypes by use of high-resolution melting-curve analysis of the hexon gene region. *Journal of Clinical Microbiology*, 47(2), 311–321.
- Stewart, P. A., & McLean, W. P., (2005). Public opinion toward the first, second, and third generations of plant biotechnology. *In Vitro Cellular & Developmental Biology-Plant*, 41(6), 718–724.
- Stoddard, B. L., (2005). Homing endonuclease structure and function. *Quarterly Reviews of Biophysics*, 38(1), 49.
- Sulakvelidze, A., Alavidze, Z., & Morris, J. G., (2001). Bacteriophage therapy. *Antimicrobial Agents and Chemotherapy*, 45(3), 649–659.
- Tahbaz, N., Subedi, S., & Weinfeld, M., (2012). Role of polynucleotide kinase/phosphatase in mitochondrial DNA repair. *Nucleic Acids Research*, 40(8), 3484–3495.
- Takeuchi, R., Lambert, A. R., Mak, A. N. S., Jacoby, K., Dickson, R. J., Gloor, G. B., Scharenberg, A. M., et al., (2011). Tapping natural reservoirs of homing endonucleases for targeted gene modification. *Proceedings of the National Academy of Sciences*, 108(32), 13077–13082.
- Tester, M., & Langridge, P., (2010). Breeding technologies to increase crop production in a changing world. *Science*, 327(5967), 818–822.
- Thomas, J. J., (2008). *Innovation in India and China: Challenges and Prospects in Pharmaceuticals and Biotechnology*. Citeseer.
- Valdez-Cruz, N. A., Caspeta, L., Pérez, N. O., Ramírez, O. T., & Trujillo-Roldán, M. A., (2010). Production of recombinant proteins in *E. coli* by the heat-inducible expression system based on the phage lambda pL and/or pR promoters. *Microbial Cell Factories*, 9(1), 1–16.
- Van, D. S., Molloy, P., Varinli, H., Morrison, J., & Muhlhauser, B., (2015). Epigenetics and human obesity. *International Journal of Obesity*, 39(1), 85–97.
- Verma, I. M., (1977). The reverse transcriptase. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 473(1), 1–38.
- Verma, S., Dutta, T., Mahadevan, S., Nalini, P., Basu, D., Biswal, N., Ramesh, A., et al., (2016). A randomized study of very low-dose factor VIII prophylaxis in severe haemophilia—A success story from a resource-limited country. *Haemophilia*, 22(3), 342–348.
- Wallace, R. B., Johnson, M. J., Suggs, S. V., Ken-ichi, M., Bhatt, R., & Keiichi, I., (1981). A set of synthetic oligodeoxyribonucleotide primers for DNA sequencing in the plasmid vector pBR322. *Gene*, 16(1–3), 21–26.
- Wilson, M. H., Coates, C. J., & George, Jr. A. L., (2007). PiggyBac transposon-mediated gene transfer in human cells. *Molecular Therapy*, 15(1), 139–145.
- Wong, D. W., (2006). Cloning vectors for introducing genes into host cells. *The ABCs of Gene Cloning*, 93–124.
- Yamaguchi, T., & Blumwald, E., (2005). Developing salt-tolerant crop plants: Challenges and opportunities. *Trends in Plant Science*, 10(12), 615–620.
- Yi, D., (2008). Cancer, viruses, and mass migration: Paul Berg's venture into eukaryotic biology and the advent of recombinant DNA research and technology, 1967–1980. *Journal of the History of Biology*, 41(4), 589–636.

- Yuan, Q., Zhang, Y., Chen, T., Lu, D., Zhao, Z., Zhang, X., Li, Z., et al., (2012). Photon-manipulated drug release from a mesoporous nanocontainer controlled by azobenzene-modified nucleic acid. *ACS Nano*, 6(7), 6337–6344.
- Zachariah, S., & Pappachen, L., (2009). A study of genetic engineering techniques in biotechnology-based pharmaceuticals. *The Internet Journal of Nanotechnology*, 3(1).
- Zeitlin, L., Pettitt, J., Scully, C., Bohorova, N., Kim, D., Pauly, M., Hiatt, A., et al., (2011). Enhanced potency of a fucose-free monoclonal antibody being developed as an Ebola virus immunoprotectant. *Proceedings of the National Academy of Sciences*, 108(51), 20690–20694.

CHAPTER 4

Introduction of Recombinant DNA into Host Cells

RIDA SALEEM, SITARA NASAR, and SAIMA IFTIKHAR

School of Biological Sciences, University of the Punjab, Lahore, Pakistan

ABSTRACT

The successful synthesis of recombinant DNA is followed by its introduction into a host cell. A recombinant DNA outside the host cell is worthless. To be meaningful, it has to transcribe and translate which requires a host cell. The host provides machinery for the replication and expression of recombinant DNA, which allows its cloning and expression respectively. Therefore, scientists introduce recombinant DNA into the host cells to achieve their goals. In this chapter, we would be discussing different methods that are commonly used for introducing recombinant DNA into the host cell.

4.1 INTRODUCTION

Processes of introducing and incorporating exogenous DNA into host cells with a stable genetic transformation have a lot of significance in cellular and molecular biology. These goals of introducing foreign DNA into the host cells can be achieved by using several methods such as transformation, transduction, and transfection, which utilizes the incorporation of DNA into the host cell by a host-specific vector or through vector-free

transformation, which primarily depends on the mechanism used for the experiment and objectives of the study. Several chemical, biological, and physical methodologies can be used depending upon the experimental needs and host cell nature. An ideal approach is to choose a method that have high transformation efficiency, low host cell toxicity, and minimum effect on the normal physiology of the transformed cells.

4.2 DIFFERENCE BETWEEN TRANSFORMATION, TRANSFECTION, AND TRANSDUCTION

Transformation, transduction, and transfection are some commonly used terminologies that are opted when manipulated DNA is introduced into host cells. The main differences between these terminologies are given below to evade misperception about them.

When recombinant DNA (rDNA) is introduced into prokaryotic competent cells by using non-viral gene transfer methods, it is usually termed as transformation. This process is not likely to happen naturally in an environment, but scientists have been harnessing the method in laboratories for so long to get eugenic aims. In this method, cells are made competent to intake foreign DNA by using several chemicals and physical means. Later on, transformed cells can be distinguished from non-transformed ones by any selectable marker gene, e.g., antibiotic-resistant genes, available in the vector.

Transfection, on the other hand, is a term applied when transferring a foreign DNA molecule into eukaryotic cells, by using non-viral means. Unlike transformation where the host cell is of prokaryotic origin, transfection deals with the uptake of manipulated DNA into eukaryotic cells. Here, eukaryotic cell membrane is made permeable by using chemical and physical methods for foreign DNA incorporation. However, methodologies used in this case are different from those of transformation. Cationic lipids, micelles, lasers, or even particle guns are some common methods used for lab-cultured eukaryotic cell transformation.

Despite transformation and transfection, transduction is another prominent terminology used for rDNA transfer. Transduction is a process of transferring modified DNA to cells, either eukaryotic or prokaryotic, by using virus-mediated DNA transfer means. To achieve this purpose, many types of virus entities, such as bacteriophages for bacteria and adenoviruses for humans, are utilized according to the objectives and requirements of the experiments.

4.3 COMMONLY USED APPROACHES TO TRANSFER RECOMBINANT DNA (RDNA) INTO HOST CELLS

There is no solitary recommended method that can be utilized for all cells or for all transformation procedures. There are many approaches, including chemical, physical, or biological-based means that can be used to achieve these objectives. Anyhow, an ideal approach is one that fulfills the experimental requirements and are acquainted to the host cell type. For instance, those methodologies that have high efficacy of transformation, and have reduced rate of toxicity for the corresponding host cells, as well as, have very little effect on host normal physiology are recommended (Kim and Eberwine, 2010). General methods applied for rDNA transfer, whether it is transformation, transfection or transduction, are elucidated in Table 4.1. Different methods for transformation are classified on the basis of versatility in their modes, and means of delivering the recombinant gene of interest into different host cells.

TABLE 4.1 Different Methods to Transfer rDNA into Host Cells

Different Methods to Transfer rDNA into Host Cells		
Chemical Approaches		
Methods	Transformed Hosts	References
CaCl ₂ treatment	Bacteria	Li et al. (2010)
Calcium phosphate mediated co-precipitation of DNA	Plants	Ardekani et al. (2014)
–	Animals	Roy et al. (2003)
DEAE-dextran approach	Animals	Pazzagli et al. (1992)
Cationic liposomes approach	Animals	Mounkes et al. (1998)
Cationic polymer approach	Animals	Ramsay et al. (2000)
Biological Approaches		
Viral vector-based rDNA delivery	Invertebrates	Jarvis et al. (1996)
–	Plants	Lico et al. (2008)
–	Animals	Vile and Russell (1995)
<i>Agrobacterium tumefaciens</i> mediated rDNA delivery	Plants	Aggarwal et al. (2011)
–	Fungi	Sharma et al. (2006)
–	Animals	Bulgakov et al. (2006)
Physical Approaches		
Particle bombardment method	Bacteria	Yoshida and Sato (2009)

TABLE 4.1 (Continued)

Different Methods to Transfer rDNA into Host Cells		
Chemical Approaches		
Methods	Transformed Hosts	References
—	Plant	Klein (1992)
—	Animals	Johnston (1990)
Microinjection	Plant	Neuhaus and Spangenberg (1990)
—	Animals	Gordon et al. (1980)
Electroporation	Bacteria	Calvin and Hanawalt (1988)
—	Plant	Fromm et al. (1986)
—	Animals	Chu et al. (1987)
Laserfection	Plant	Guo et al. (1995)

4.3.1 CHEMICAL APPROACHES TO TRANSFER RECOMBINANT DNA (RDNA) INTO HOST

rDNA can be transferred into a variety of host cells ranging from prokaryotic cells to eukaryotic cells by using non-viral chemical-based means. These chemical approaches are not only easy to manipulate, but they also have reduced the problems of cell toxicity effectively. There are several suggested methods for successful delivery of gene of interest into respective host cells, elucidated in Figure 4.1.

4.3.1.1 CALCIUM SALT TREATMENT

The use of calcium salt for transformation of foreign DNA involve different mechanisms for prokaryotic and eukaryotic cells. In prokaryotic cells, the mechanism of action involves pre-treatment of bacterial host cells with calcium chloride to generate pores in their cell membrane, as a result, competent cells are produced. Competent cells have the potential to uptake rDNA easily and effectively. In the case of plants, protoplast is generated for plant cells by removing plant cell wall, and then salt treatments allow plant cells to take DNA into them; however, transforming foreign DNA into animal cells involves an entirely different methodology.

Mammalian cells are nowadays in demand for being a suitable potential host to produce recombinant protein as compared to prokaryotic cells because they can perform post-translational modifications of protein.

Transformation of foreign DNA into mammalian cell cultures can be done efficiently using chemical means. This was introduced when researchers formed calcium phosphate DNA co-precipitates by mixing DNA into calcium chloride suspension and dispersing it over human KB cell line, subsequently achieving positive results (Graham and van der Eb, 1973). The principle behind this methodology is that calcium phosphate-based co-precipitates can easily be obtained by mixing purified rDNA into

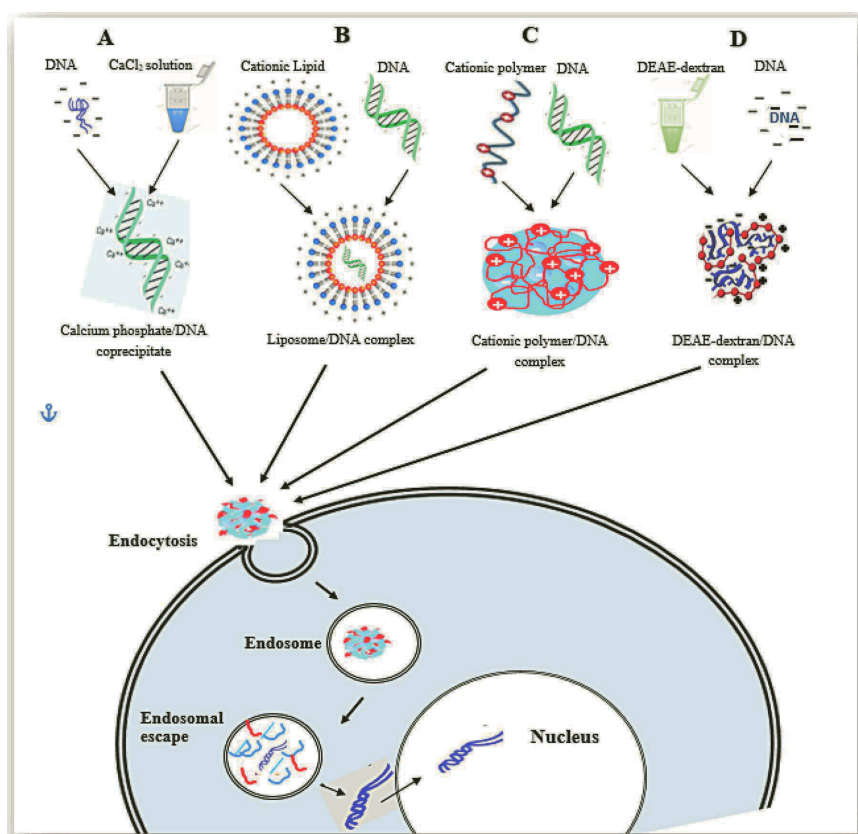


FIGURE 4.1 Schematic representation of chemical facilitated DNA delivery to host cell.

Note: (A) Calcium phosphate co-precipitation mediated DNA delivery; (B) cationic lipid-based approach; (C) cationic polymer mediated delivery; and (D) DEAE-dextran mediated approach. All aforementioned chemical methods follow the same mechanism commencing from entry into the cell through endocytosis to endosomal discharge, and subsequently translocation of DNA towards the nucleus.

calcium chloride and sodium phosphate solution. These co-precipitates of calcium phosphate are dispersed onto the respective mammalian cell surface subsequently leads towards the entry into host cells via endocytosis. Later, endosomal escape is mediated by calcium ions and subsequently DNA discharged into cytosol.

Calcium phosphate-based micro-precipitates loaded with DNA have tremendous importance in molecular biology for efficient transfer of DNA into host cells due to its several kinds of advantages. One of its important benefits over other techniques is, it avoids detrimental effects of viral vector-based DNA delivery. Furthermore, its efficiency is far greater than other non-viral DNA delivery systems, which can be enhanced by using fine and smallest calcium phosphate precipitates. Likewise, they provide an effective endosomal discharge in the cytoplasm and provide a stabilized translocation towards the nucleus.

Despite of advantages there are some drawbacks as well. These have almost 15% less DNA transfer efficiency as compared to viral vectors. Precipitates size is another worse drawback of this approach because calcium phosphate precipitates of approximately 300 nm are considered compatible to use for mammalian cell lines and having low cell toxicity. But this size is too large to encapsulate and carry DNA into it, consequently allows slow DNA transfer to host cells. However, synthetic viral sized calcium phosphate nano-precipitates having size less than 100 nm are now in experimental trials (Maitra, 2005).

4.3.1.2 DEAE-DEXTRAN TRANSFECTION

Diethylaminoethyl-dextran (DEAE-dextran) polymer is a carbohydrate ploy-cationic derivative and is an ancient method used for non-viral recombinant plasmid DNA transfer into mammalian host cells. It is considered the first ever reported chemical used for DNA transformation that was reported in 1965 to escalate viral infectivity of cells (Vaheri and Pagano, 1965). Due to the highly positive charge on DEAE, it can bind to negatively charged DNA due to ionic interactions but only suitable for transient DNA delivery into host cells. The mechanism of action is that a solution comprising of DEAE-dextran polymer and recombinant plasmid DNA moiety is used to incubate the cultured host cells. In the growth media chloroquine is added to give stabilization to recombinant plasmid

DNA since chloroquine can degrade lysosomal hydrolases. To escalate the permeability of host cells dimethyl sulfoxide (DMSO) shock is also given to the media (Sussman and Milman, 1984). It helps in efficient endocytosis, followed by endosomal discharge inside cytoplasm due to a drop in pH and translocation of DNA to the nucleus for gene expression (Gulick, 2001).

The major advantages of using this approach are, its high-speed delivery, easy manipulation, and low cost. Some of the drawbacks of using this carbohydrate derivative for DNA delivery is that it is only limited to a short range of cell types and have very low efficiency as compared to all other DNA delivery methods. Moreover, it does not generate stable cell line and is suitable for transient transfection (Lalani and Misra, 2011). Other disadvantages are limited to its optimization since concentration of DEAE-dextran, incubation times, and different types of helpers, e.g., chloroquine and DMSO optimization are some factors that influence efficient plasmid DNA delivery by using this approach (Kumar et al., 2018).

4.3.1.3 CATIONIC LIPID MEDIATED RECOMBINANT DNA (RDNA) TRANSFER INTO HOST CELL

Cationic lipid-based rDNA transfer to host cells is one of the significant and popular approaches utilized in rDNA technology. Lipoplex is a term coined for complexes composed of cationic liposomes and DNA that is widely used in biotechnology and molecular biology labs for non-viral DNA transfer, usually to animal cell cultures. Liposomes are normally comprised of two components, first one is the cationic lipid and other one is neutral lipid (also renowned for helper lipid). Cationic lipid has at one side a cationic head is attached to a hydrophobic tail with the aid of a linker while the other end is attached to neutral lipids, which are cholesterol and dioleoylphosphatidyl ethanolamine (DOPE). DOPE helps cationic lipid for the formulation of a bilayer once they get attached to DNA moiety and also helps in attachment with the plasma membrane of host cells. Cholesterol is seemed to assist in stabilization of complex in physiological medium subsequently allows the complex to reach at the target site without degrading the DNA (Hafez et al., 2001). For the very first time, this method was used by entrapping Simian virus 40 (SV40) viral DNA into 0.4 μm diameter liposomes and introducing liposomes-enclosed

viral DNA complex into monkey cell lines (Fraley et al., 1980). The principle behind this approach is that a net positive charge over the surface of cationic lipid/DNA complex facilitates efficient adsorption of lipoplexes (having rDNA or gene of interest) through the anionic plasma membrane of host cells via electrostatic interactions (Ilarduya et al., 2010).

The mechanism of cationic lipid/DNA complex formation is still under investigation. However, the closely accepted hypothesis in this regard is that being charged particles both cationic lipid and DNA are driven by electrostatic interactions and DNA entrapped inside the liposomes forming complex lipoplex. This complex attaches to the plasma membrane of the host cell by ionic interactions due to positive lipoplex and negative cell membrane, which subsequently allows endocytosis of lipoplex. This results in double-layered micelle formation as endosome. Endosome undergoes maturation into lysosomes and pH of lipid compartments of endosomes changes from pH 7 to 5.5, which leads to the rupture of the outer lipid bilayers (Ilarduya et al., 2010). Meanwhile, the DNA releases into the cytosol and ultimately translocate towards the nucleus where it carries out gene expression.

Like all other rDNA technology, this cationic lipid-mediated DNA delivery system has some pros and cons. One of its biggest advantages is that this approach is efficiently applicable on a wide range of animal and many other eukaryotic cell lines and for effective delivery of any size of DNA. But its core disadvantage is that due to its specificity for cell lines and for culture conditions, there are a lot of factors that have to be optimized to carry the approach. These factors include the nature of cationic lipid reagent, interaction of cationic lipid with rDNA, interaction of lipoplex with cell membrane of different cell lines and release of DNA inside the cytoplasm of respective cell line.

Life Technologies™ offers a wide variety of cationic lipid reagents for different types of cell type and according to different culture conditions, such as Lipofectamine® 3000 reagent (Cohen, 2009) that is one of the popular cationic lipid reagents used in laboratories for rDNA transfer.

4.3.1.4 RECOMBINANT DNA (RDNA) DELIVERY BY CATIONIC POLYMERS-BASED APPROACHES

Cationic polymers are different from cationic lipids both structurally and functionally however, they share some common features as well. Cationic

polymers, unlike cationic lipids lack hydrophobic region in their structures and are completely soluble in polar aqueous media (Lv et al., 2006). Functionally, cationic polymers are more effective and efficient in delivering rDNA into host cells. In biochemical studies, they have got tremendous importance for non-viral chemical vector-based DNA delivery. The principle behind their mechanism of action is identical to other chemical-based non-viral vectors that they form complex with DNA on the basis of electrostatic interactions and adsorbed to the cell membrane of host cells, followed by endocytosis and endosomal discharge inside the cytosol of target host cell rDNA is released and translocated towards the nucleus for gene expression.

Cationic polymers are classified on the basis of their nature, origin, and structures as peptide-based cationic polymers, branched or unbranched cationic polymers, natural cationic polymers, and polysaccharide-based cationic polymers. Significant examples of peptide-based cationic polymers are polylysine and polyornithine, as well as, polysaccharide-based cationic polymers are cyclodextrin. Despite of peptide-based cationic polymers, synthetic branched, and unbranched polymers are also commercially available, including polyethyleneimine and polybrene. Natural polymers such as collagen is also significant. These are some important commercial cationic polymers available for efficient DNA delivery into host cells (Ramsay et al., 2000).

4.3.2 BIOLOGICAL METHODS TO TRANSFER RECOMBINANT DNA (RDNA) INTO HOST

Significance of biological methods in gene delivery approaches cannot be neglected. Despite the discovery of several chemicals and physical methodologies, biological methods, for instance, viral vectors-based methods and agrobacterium-based methods, are of huge importance because of their high efficiency rate of nucleic acid transfer.

4.3.2.1 VIRAL VECTOR-MEDIATED RECOMBINANT DNA (RDNA) DELIVERY INTO HOST CELLS

Viral vector-based DNA transfer into host cell is termed as transduction. Viruses can be used as a successful vehicle for rDNA delivery into host cells, where the target purpose is the expression of recombinant protein.

Since viruses are host specific by nature, this is the reason for the synthesis of a wide range of viral vectors at the commercial level. DNA delivery method, however, is based on the type of cells and transfection efficiency. Among them most commonly used DNA viral vectors according to their host specificity are bacteriophages for bacterial hosts, retroviruses-based vectors for mammalian cells (Vile and Russell, 1995), baculoviruses specific for invertebrates (Jarvis et al., 1996) and plant-based viral vectors (Lico et al., 2008). However, herpes viral-based vectors, adenoviral vectors, and retroviral vectors are used in more than 70% of biochemical rDNA delivery research methods, among them most prominent is retro-virus (Walther and Stein, 2000).

The mechanism of action involves a two-step process; first is a synthesis of recombinant viruses having genes of interest in it, followed by its amplification in packaging cell lines. Then in the next step, viral particles are purified and allowed to infect their specific host cell line and to achieve rDNA expression (Warnock et al., 2006).

There are significant advantages, as well as, disadvantages of using viral vector mediated DNA delivery systems. Major advantages of using viral vectors are their great amplification rate, they are easy to manipulate, their escalated translational rate inside the host, and site specificity for different hosts. All these advantages make viral-based vectors a suitable choice to use it for rDNA technology (Lico et al., 2008). But they also have some disadvantages that cannot be neglected. The drawbacks include host cells immune response towards the viral antigens when recombinant viral vector enters into host cells. Host cell toxicities and expensive approach because of biosafety requirements are also other big issues of using viral vectors for DNA delivery. Some other challenges are virus packaging capacity that is very low than non-viral DNA delivery vectors and host specificity.

4.3.2.2 AGROBACTERIUM-MEDIATED RECOMBINANT DNA (RDNA) DELIVERY

Agrobacterium tumefaciens was notorious for bringing about crown gall disease in plants, specifically in the members of the rose family at the time of its discovery (Smith and Townsend, 1907). But elegant research made on its mechanism of action, enabled scientists to be used it as a tool for

effective transformation of genetic material into plant cells (Binns and Thomashow, 1988). The system got tremendous importance in molecular biology for having the potential to transfect not only the plants, but also the animal cells (Bulgakov et al., 2006) and fungal species (Sharma et al., 2006).

Natural mechanism of infection of *A. tumefaciens* in plants can be elucidated as the soil bacteria transfers its specific T-DNA of the tumor-inducing plasmid (Ti-plasmid) into the plant cell, where it integrates with the host genome (Opabode, 2006). Transformation of the T-DNA into the cell is regulated by the virulence genes of the *A. tumefaciens* transformed genetic material (Bulgakov et al., 2006). T-DNA was reported to have two types of genes (Opabode, 2006); the oncogenic genes that seems to involve in tumor production, and the non-oncogenic genes encode for opines (condensation product of natural sugars and amino acids that fulfills nutritional needs of *A. tumefaciens*). To get eugenic aims, T-DNA vector is genetically engineered (GE) devouring a selectable marker gene alongside with the gene of interest or rDNA, placed between T-DNA borders. Subsequently, by using the same principle of natural mechanism of infection, plant cells are transformed (Sparks et al., 2014).

Major advantages are the successful transformation of several crops, cereals (including maize, wheat, and barley) with potent tissue culture responses. rDNA is transferred with minimum re-arrangements, as well as, well-demarcated ends with high fertility of the host cells.

4.3.3 PHYSICAL APPROACHES TO TRANSFER RECOMBINANT DNA (RDNA) INTO HOST

Physical methods have also acquired hold of the advanced level in rDNA technology. Some of the physical approaches are elaborated below that provide chemical free and viral free rDNA delivery and avoid cytotoxicity level to much extent.

4.3.3.1 TRANSFER OF DNA TO HOST VIA ELECTROPORATION

Electroporation is an important physical approach to transfer rDNA into host cells. It involves the generation of transient pores in host cells plasma

membrane, by applying electric pulse, for intake of genetic material with greater efficiency in rDNA technology. This methodology can be effectively applied for both prokaryotic and eukaryotic cells, frequently used for plant and algae, while less commonly applied for bacteria and mammalian cells. Prokaryotic cells due to its very small size as compared to eukaryotic cells require high voltage and greater time to induce pores in the outer membrane (Calvin and Hanawalt, 1988). Pulse generator apparatus is used to generate an electric pulse of required voltage and vital time exposure according to cell types and optimized conditions for electroporation.

Working principle is that host cells along the rDNA are suspended in a conductive solution inside the pulse generator. Voltage and optimal conditions are set according to the nature of the host cell. Electric current is generated for very short interval, and transient pores are generated due to disturbance of phospholipid bilayers of the membranes. Rise in electric field across the cell membranes allows easy and fast intake of rDNA into the respective host cells (Shigekawa and Dower, 1988).

The main advantage of using this method is that it is user friendly and rapid to carry out. One of its drawbacks is that its efficiency is low for prokaryotic cells due to their small sizes as *Escherichia coli* (*E. coli*) cells transformation efficiency is 10^2 – 10^4 per microgram of rDNA while, transformation efficiency with calcium chloride treatment for *E. coli* is 10^5 – 10^6 per microgram of rDNA (Calvin and Hanawalt, 1988). Another disadvantage is that high voltage can cause cell death that can be avoided by using contemporary electroporation systems such as Invitrogen™ Neon™ electroporation system that offers equal distribution of electric pulse inside conductive solution and decrease mortality rate of host cells (Shigekawa and Dower, 1988).

4.3.3.2 BIOLISTIC PARTICLE DELIVERY OF DNA INTO HOST CELLS

Particle bombardment is a contemporary physical technique for direct transfer of rDNA into host cells. Alternative terms most frequently used for the respective methodology are particle-gun method, micro-projectile approach, particle bombardment method, biolistic particle delivery, bio-blaster method and the gene-gun approach (Sanford, 1988). The working principle behind this approach is that rDNA is first modified in the form of particles, ranging from 0.5–5.0 μm , by using metallic powder mostly

gold or tungsten. Sizes of the particles are optimized according to type and permeability of the cell. A small amount of this slurry is placed into the particle gun device and bombardment is allowed from a tiny hole with high velocity. At this step, conditions for various cell lines can be varied according to variable optimization (Johnston, 1990). Optimization is achieved by varying several physical factors such as momentum of particles (DNA coated with metal powder), distance between particle gun and host cells, and rDNA concentration, as well as, several biological factors such as cell type and physiological conditions of the cells (Klein, 1992).

Micro projectile rDNA delivery can be applied to microbe, plant, or animal cells (Klein, 1992). But this approach is the most effective for the plant cells, where without removing the cell wall rDNA can be injected into intact plant cells via biolistic gun (Klein, 1990). Furthermore, it is the only source which can be used for delivering DNA to chloroplasts or mitochondria. Cell types of maize, soybean, and rice have been testified and reported until now that undergo stable transformation via particle gun method. As far as, microbial hosts are concerned, the particle gun method has not got much importance to them. One of the basic reasons is that this process is expensive, and alternative inexpensive physical, chemical, and biological procedures are available for their transformation. Anyhow, there are some drug-producing *Streptomyces* microbial species that are commercially viable strains, but there is not yet any proper transformation system reported for them. So, particle gun has got much importance for such type of microbial cell lines in recent years. Similarly, mammalian cells transfection with biolistic approach is also a recent event (Johnston, 1990). The most noteworthy feature here is the successful biolistic rDNA delivery to animal cells, as well as, to several organs such as the liver, spleen, and skin in live animals. Despite of using traditional particle bombardment methods, due to complex conditions, somatic cells of live animals are transfected by using modern technology, including electric discharge and using instruments based on gunpowder (Williams et al., 1991).

The main advantages of this approach are that this is the only way for rDNA delivery to organelles of plant cells, organs of mammalian cells, somatic cells of live animals and viral-based vector-free delivery. Regardless of advantages, there are some drawbacks as well. They include the requirement of special instrumentation and high expenditures.

4.3.3.3 *DIRECT MICROINJECTION OF GENE OF INTEREST INTO HOST CELLS*

Direct microinjection involves the procedures used to transfer rDNA directly into host cell (one cell at a time) with the help of glass pipettes. The technique of direct micro-injection of nucleic acid into host cells with the help of micro-pipettes was used successfully for the very first time by Graessmann (1968). This technique is most frequently applied for plant cells, but also a number of successes have been achieved for gene delivery into microbes and animal cells. All the systems involve common procedure for the DNA microinjection although some variables are always there according to experimental conditions. The basic mechanism is adherence of the target cells to some surfaces that provide proper anchoring for microinjection that will aid in proper orientation and precise microinjection. Agarose and sodium alginate-based immobilization surfaces are reported anchoring surfaces suitable for microinjection (Neuhaus and Spangenberg, 1990). Then DNA is injected into the cells with the help of micro-needles. The efficiency of using this approach is extremely high but the number of transformed cells is very low, i.e., usually 100–200 cells per hour (Neuhaus and Spangenberg, 1990). Besides dealing with the protoplast of plant cells, intact cells along with cell wall has been microinjected as well. This system involves the use of highly competent single cells for microinjection (Nomura and Komamine, 1985).

Unlike other biological, chemical, or physical gene delivery approaches, microinjection is the only method where rDNA is the single variable parameter.

The main advantages of this technique are its high-efficiency rate and its suitability for transduction-challenged cells. Moreover, cytotoxicity is very low and highly precise concentration of rDNA can be injected. But it is extremely labor intensive because the cell mortality ratio is very high, high skill demanding and a very small number of cells can be transformed, that are its worse limitations.

4.3.3.4 *LASER-MEDIATED TRANSFECTION*

Laser-mediated transfection also termed as photo-transfection or laser-fection. Collimated laser beams have got tremendous importance for opto-poration of single cell due to low divergence of laser beam in past

few years. However, wavelength range of UV laser pulse between 337 to 355 nm, as well as, wavelength of 532 nm for visible laser pulse have been recorded the most suitable for opto-poration of cell membrane and loading exogenous DNA into the cell (Schneckenburger et al., 2002). Laser-mediated transfection is defined as the intake of nucleic acid by the host cells through generating transient pores in the cell membrane with the help of a laser pulse. When a laser pulse generates pores in the membrane, then osmotic pressure difference between cytosol and media containing nucleic acid allows DNA to enter into the cells. To establish a precise osmotic pressure, plant cells are pre-treated in a hypertonic solution to pull some water out of the host plant cells. These cells are put into a hypotonic solution having an rDNA molecule that will generate a precise osmotic difference. Pores are induced straightaway with the laser beam of optimized wavelength which facilitates DNA uptake by the cells. Various fluorescent markers can be used to confirm the entry of exogenous nucleic acid into the cell (Guo et al., 1995).

With photo-transfection suitable pores can be induce at desirable cell surface that is its one prime advantage. But its limitations are due to be an expensive procedure as costly laser microscopic systems are required.

4.4 FUTURE PERSPECTIVES

Innovative and advanced approaches regarding rDNA delivery into host cells are evolving relentlessly. Researches are endeavoring to search more contemporary ways to transfer DNA into the host with minimum cytotoxicity and maximum efficiency. The available approaches, including biological, chemical, and physical, have certain limitations and drawbacks. These downsides of the experiments are sometimes might be about the cargo size, inefficiency of the delivery system, high-cost consumption, cytotoxicity issues, or difficult manipulations. These limitations are a big halt in carrying out the experiment efficiently. These are some basic experimental issues that require dire attention and improvement in future technologies. Future progress cannot be limited only to get efficiency in the available approaches for transformation, but the future objectives can be extended to get sub-cellular DNA transfer (that can be either dendrites or can be the mitochondria and chloroplast like regions), as well as, to transform a whole individual.

To conclude, despite of a lot of advances in the field of rDNA transfer in cellular and molecular biology, more advanced options are still in demand.

KEYWORDS

- **cationic lipids**
- **diethylaminoethyl-dextran**
- **dimethyl sulfoxide**
- **dioleylphosphatidyl ethanolamine**
- **molecular biology**
- **recombinant DNA**

REFERENCES

- Aggarwal, D., Kumar, A., & Reddy, M. S., (2011). Agrobacterium tumefaciens mediated genetic transformation of selected elite clone (s) of *Eucalyptus tereticornis*. *Acta Physiologiae Plantarum*, 33(5), 1603–1611.
- Ardekani, M. R. S., Abidin, M. Z., Nasrullah, N. A. Z. I. M. A., & Samim, M., (2014). Calcium phosphate nanoparticles a novel non-viral gene delivery system for genetic transformation of tobacco. *Int. J. Pharm. Pharm. Sci.*, 6(6), 605–609.
- Binns, A. N., & Thomashow, M. F., (1988). Cell biology of agrobacterium infection and transformation of plants. *Annual Reviews in Microbiology*, 42(1), 575–606.
- Bulgakov, V. P., Kiselev, K. V., Yakovlev, K. V., Zhuravlev, Y. N., Gontcharov, A. A., & Odintsova, N. A., (2006). Agrobacterium-mediated transformation of sea urchin embryos. *Biotechnology Journal*, 1(4), 454–461. doi: 10.1002/biot.200500045.
- Calvin, N. M., & Hanawalt, P. C., (1988). High-efficiency transformation of bacterial cells by electroporation. *Journal of Bacteriology*, 170(6), 2796–2801. doi: 10.1128/jb.170.6.2796-2801.1988.
- Celis, J. E., (1984). Microinjection of somatic cells with micropipettes: Comparison with other transfer techniques. *Biochemical Journal*, 223(2), 281–291. doi: 10.1042/bj2230281.
- Chu, G., Hayakawa, H., & Berg, P., (1987). Electroporation for the efficient transfection of mammalian cells with DNA. *Nucleic Acids Research*, 15(3), 1311–1326.
- Cohen, R. N., Marieke, A. E. M., Macaraeg, N., Lee, A. P., & Szoka, F. C., (2009). Quantification of plasmid DNA copies in the nucleus after lipoplex and polyplex transfection. *Journal of Controlled Release*, 135(2), 166–174.

- Fraley, R., Subramani, S., Berg, P., & Papahadjopoulos, D., (1980). Introduction of liposome-encapsulated SV40 DNA into cells. *Journal of Biological Chemistry*, 255(21), 10431–10435.
- Fromm, M. E., Taylor, L. P., & Walbot, V., (1986). Stable transformation of maize after gene transfer by electroporation. *Nature*, 319(6056), 791–793.
- Gordon, J. W., Scangos, G. A., Plotkin, D. J., Barbosa, J. A., & Ruddle, F. H., (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. *Proceedings of the National Academy of Sciences*, 77(12), 7380–7384.
- Graessmann, A., (1968). *Doctoral Dissertation*. Free University of Berlin.
- Graham, F. L., & Van, D. E. A. J., (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. *Journal of Virology*, 52(2), 456–467. [https://doi.org/10.1016/0042-6822\(73\)90341-3](https://doi.org/10.1016/0042-6822(73)90341-3).
- Gulick, T., (2001). Transfection using DEAE-dextran. *Current Protocols in Molecular Biology*. doi: 10.1002/0471142727.mb0902s40.
- Guo, Y., Liang, H., & Berns, M. W., (1995). Laser-mediated gene transfer in rice. *Physiologia Plantarum*, 93(1), 19–24. doi: 10.1034/j.1399-3054.1995.930104.x.
- Hafez, I., Maurer, N., & Cullis, P., (2001). On the mechanism whereby cationic lipids promote intracellular delivery of polynucleic acids. *Journal of Gene Therapy*, 8, 1188–1196. <https://doi.org/10.1038/sj.gt.3301506>.
- Ilarduya, C. T., Sun, Y., & Düzgüneş, N., (2010). Gene delivery by lipoplexes and polyplexes. *European Journal of Pharmaceutical Sciences*, 40(3), 159–170.
- Jarvis, D. L., Weinkauff, C., & Guarino, L. A., (1996). Immediate-early baculovirus vectors for foreign gene expression in transformed or infected insect cells. *Protein Expression and Purification*, 8(2), 191–203.
- Johnston, S. A., (1990). *Biolistic Transformation: Microbes to Mice*. *Journal of Nature*, 346(6286), 776–777. doi: 10.1038/346776a0.
- Kim, T. K., & Eberwine, J. H., (2010). Mammalian cell transfection: The present and the future. *Journal of Anal. Bioanal. Chem.*, 397, 3173–3178. <https://doi.org/10.1007/s00216-010-382-6>.
- Klein, T. M., Arentzen, R., Lewis, P. A., & Fitzpatrick-McElligott, S., (1992). Transformation of microbes, plants and animals by particle bombardment. *Journal of Nature Biotechnology*, 10(3), 286–291. doi: 10.1038/nbt0392-286.
- Klein, T. M., Goff, S. A., Roth, B. A., & Fromm, M. E. (1990). Applications of the Particle Gun in Plant Biology. *Current Plant Science and Biotechnology in Agriculture*, 56–66. doi:10.1007/978-94-009-2103-0_8.
- Kumar, P., Nagarajan, A., & Uchil, P. A., (2018). DEAE-dextran transfection. *Cold Spring Harbor Protocols*, (7). doi: 10.1101/pdb.top096263.
- Lalani, J., & Misra, A., (2011). Gene delivery using chemical methods. *CHALLENGES in Delivery of Therapeutic Genomics and Proteomics*, 127–206. Doi:10.1016/b978-0-12-384964-9.00004-9.
- Li, X., Sui, X., Zhang, Y., Sun, Y., Zhao, Y., Zhai, Y., & Wang, Q., (2010). An improved calcium chloride method preparation and transformation of competent cells. *African Journal of Biotechnology*, 9(50), 8549–8554.
- Lico, C., Chen, Q., & Santi, L., (2008). Viral vectors for production of recombinant proteins in plants. *Journal of Cellular Physiology*, 216(2), 366–377. doi: 10.1002/jcp.21423.

- lv, H., Zhang, S., Wang, B., Cui, S., & Yan, J., (2006). Toxicity of cationic lipids and cationic polymers in gene delivery. *Journal of Controlled Release*, 114(1), 100–109. doi: 10.1016/j.jconrel.2006.04.014.
- Maitra, A., (2005). Calcium phosphate nanoparticles: Second-generation nonviral vectors in gene therapy. *Journal of Expert Review of Molecular Diagnostics*, 5(6), 893–905. <https://doi.org/10.1586/14737159.5.6.893>.
- Mounkes, L. C., Zhong, W., Cipres-Palacin, G., Heath, T. D., & Debs, R. J., (1998). Proteoglycans mediate cationic liposome-DNA complex-based gene delivery *in vitro* and *in vivo*. *Journal of Biological Chemistry*, 273(40), 26164–26170.
- Neuhaus, G., & Spangenberg, G., (1990). Plant transformation by microinjection techniques. *Physiologia Plantarum*, 79(1), 213–217. doi: 10.1111/j.1399-3054.1990.tb05890.x.
- Neuhaus, G., & Spangenberg, G., (1990). Plant transformation by microinjection techniques. *Physiologia Plantarum*, 79(1), 213–217.
- Nomura, K., & Komamine, A., (1985). Identification and isolation of single cells that produce somatic embryos at a high frequency in a carrot suspension culture. *Journal of Plant Physiology*, 79(4), 988–991. doi: 10.1104/pp.79.4.988.
- Ohashi, S., Kubo, T., & Ikeda, T., (2001). Cationic polymer-mediated genetic transduction into cultured human chondro sarcoma-derived HCS-2/8 cells. *Journal of Orthopaedic Science*, 6, 75–81. <https://doi.org/10.1007/s007760170028>.
- Opabode, J. T., (2006). Agrobacterium-mediated transformation of plants: Emerging factors that influence efficiency. *Biotechnology and Molecular Biology Reviews*, 1(1), 12–20.
- Pazzagli, M., Devine, J. H., Peterson, D. O., & Baldwin, T. O., (1992). Use of bacterial and firefly luciferases as reporter genes in DEAE-dextran-mediated transfection of mammalian cells. *Analytical Biochemistry*, 204(2), 315–323.
- Ramsay, E., Hadgraft, J., Birchall, J., & Gumbleton, M., (2000). Examination of the biophysical interaction between plasmid DNA and the polycations, polylysine and polyornithine, as a basis for their differential gene transfection in-vitro. *International Journal of Pharmaceutics*, 210(1, 2), 97–107.
- Roy, I., Mitra, S., Maitra, A., & Mozumdar, S., (2003). Calcium phosphate nanoparticles as novel non-viral vectors for targeted gene delivery. *International Journal of Pharmaceutics*, 250(1), 25–33.
- Sanford, J., (1988). The biolistic process. *Trends in Biotechnology*, 6(12), 299–302. doi: 10.1016/0167-7799(88)90023-6.
- Schneckenburger, H., Hendinger, A., Sailer, R., Strauss, W. S. L., & Schmitt, M., (2002). Laser-assisted optoporation of single cells. *Journal of Biomedical Optics*, 7(3), 410. doi: 10.1117/1.1485758.
- Sharma, K. K., Gupta, S., & Kuhad, R. C., (2006). Agrobacterium-mediated delivery of marker genes to *Phanerochaete chrysosporium* mycelial pellets: A model transformation system for white-rot fungi. *Biotechnology and Applied Biochemistry*, 43(3), 181–186.
- Shigekawa, K., & Dower, W. J., (1988). Electroporation of eukaryotes and prokaryotes: A general approach to the introduction of macromolecules into cells. *Journal of BioTechniques*, 6(8), 742–751.
- Smith, E. F., & Townsend, C. O., (1907). A plant-tumor of bacterial origin. *Science*, 25(643), 671–673.

- Sparks, C. A., Doherty, A., & Jones, H. D., (2014). Genetic transformation of wheat via agrobacterium-mediated DNA delivery. In: *Cereal Genomics* (pp. 235–250). Humana Press, Totowa, NJ.
- Sussman, D. J., & Milman, G., (1984). Short-term, high-efficiency expression of transfected DNA. *Molecular and Cellular Biology*, 4(8), 1641–1643. doi: 10.1128/mcb.4.8.1641.
- Vaheri, A., & Pagano, J. S., (1965). Infectious poliovirus RNA: A sensitive method of assay. *Journal of Virology*, 27(3), 434–436. doi:10.1016/0042-6822(65)90126-1.
- Vile, R. G., & Russell, S. J., (1995). Retroviruses as vectors. *British Medical Bulletin*, 51(1), 12–30.
- Walther, W., & Stein, U., (2000). Viral vectors for gene transfer. *Journal of Drugs*, 60(2), 249–271. <https://doi.org/10.2165/00003495-200060020-00002>.
- Warnock, J. N., Merten, O. W., & Al-Rubeai, M., (2006). Cell culture processes for the production of viral vectors for gene therapy Purposes. *Journal of Cytotechnology*, 50, 141–162. doi:10.1007/s10616-005-5507-z.
- Williams, R. S., Johnston, S. A., Reidy, M., DeVit, M. J., McElligott, S. G., & Sanford, J. C., (1991). Introduction of foreign genes into tissues of living mice by DNA-coated microprojectiles. *Proceedings of the National Academy of Sciences*, 88(7), 2726–2730. doi: 10.1073/pnas.88.7.2726.
- Yoshida, N., & Sato, M., (2009). Plasmid uptake by bacteria: A comparison of methods and efficiencies. *Applied Microbiology and Biotechnology*, 83(5), 791–798.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

CHAPTER 5

Linking of Desired Gene with DNA Vector/Gene Cloning Vector

SITARA NASAR AND SAIMA IFTIKHAR

School of Biological Sciences, University of the Punjab, Lahore, Pakistan

ABSTRACT

Recombinant DNA (rDNA) technology is a key for making advancements in the biotechnological industry. This enables the scientists to fuse together genetic materials of two completely different origins, to obtain specific genetic variations which are meaningful to human development. This technology has a diverse application in various life sciences such as science, agriculture, and pharmaceuticals.

In this chapter we will discuss about the very first step in rDNA technology, i.e., insertion of a desired gene into a cloning vector. It is a sequential protocol which utilizes various steps and each step requires a set of biological or chemical molecules to link a desired gene in the cloning vector. Therefore, we will first introduce those molecules followed by their usage in rDNA technology.

5.1 CLONING VECTORS

Cloning vectors provide a backbone structure to carry a target gene or any nucleotide sequence. These vectors can only amplify the magnitude of inserted gene, and do not have the potential to take part in gene expression or protein

synthesis. Expression studies, however, are carried out using another type of vectors, i.e., expression vectors which are beyond the scope of this chapter. In this chapter, we will only focus on cloning vectors (Carter and Shieh, 2010).

For linking a gene with a cloning vector, we first need a gene of interest and a cloning vector. Vectors are self-replicating, double stranded circular DNA molecules that act as a vehicle, which carry the gene of interest and transport it into a suitable host to generate multiple copies of the target gene. There are various types of cloning vector depending on the size of the insert they could carry, their copy number, and type of cloning method one is using. For a cloning vector to be used in recombinant DNA (rDNA) technology, it should have a specific set of characteristics.

5.1.1 ESSENTIAL FEATURES OF A CLONING VECTOR

5.1.1.1 ORIGIN OF REPLICATION (ORI)

DNA replication is a biological process of generating two identical DNA molecules from a single original DNA molecule. Parent DNA strand act as a template and where elements of replication machinery interact to complete a successful replication cycle. As theorized by Watson and Crick, DNA replication follows the semi-conservative replication model, where each daughter molecule carries one strand from the parent molecule.

To self-replicate itself, a cloning vector must have a particular DNA sequence called the origin of replication (Ori), a point where the replication of the vector starts to generate daughter molecules that are identical to the parent molecule (Figure 5.1). Although the organizational structure of Ori as well as its recognition varies from specie to specie, there are three essential structural features of an Ori. Firstly, it must have a site to welcome initiation and auxiliary proteins to start the replication process. Secondly, it is usually rich in adenine (A) and thymine (T) nucleotides, as they are bonded together by double hydrogen bond which is easy to separate instead of triple bond between guanine (G) and cytosine (C) base-pairings, which is a relatively stronger bond to break. Lastly, an Ori point have structural sites that regulate replication event (Howe, 2018).

There is a lot of diversity in Ori, however cloning vectors are mostly of bacterial origin, i.e., ColE1, pMB1, 15A, R6K, and pSC101 and their derivatives. Each Ori sequence has its own unique structural and functional characteristics, i.e., some have the ability to produce various copies whereas, others are tightly regulated and generate only a few copies. This

regulation is generally referred as stringent or relaxed based upon positively regulatory molecule, i.e., protein or RNA, respectively. Stringent regulation requires initiation proteins, which are synthesized by the host cell whenever replication process is required. On the other hand, relaxed regulation does not require host initiation proteins and only require machinery for elongation and termination (Del Solar et al., 1998).

Ori can be manipulated to fulfill desire needs, e.g., pBR322. It is a plasmid and its Ori region allows it to maintain approx. 20–50 copies of plasmid per cell whereas introduction of only a single point-mutations in pBR322 at 3,075 positions, where a G is substituted by T, gives rise to another plasmid named pHC, a cloning vector which produces up to 100 copies per cell (Boros et al., 1984). Similarly, a cloning vector might be manipulated to carry more than one Ori, e.g., P1 cloning vector. It is a bacteriophage that carries two Ori regions: one is of plasmid origin that regulates replication of plasmid DNA and maintains 1 copy per host chromosome, whereas, the other one is of viral origin. This lytic Ori amplifies the copy number of vector to approx. 20 copies per cell upon induction (Karcher, 1995).

To summarize this discussion, Ori is an essential sequence in a cloning vector, and scientists with deep knowledge about its sequence stability and functionality under given circumstances modify these sequences in order to achieve their target goals. Cloning vectors might be circular or linear in its structure, however most widely used vectors are circular in nature, that is why in this chapter, we used circular illustrations to represent cloning vectors. Figure 5.1 depicts an Ori region within a cloning vector.

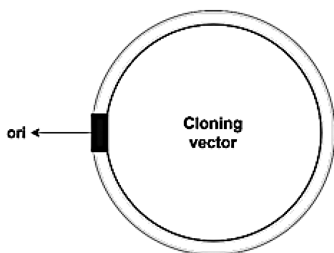


FIGURE 5.1 Origin of replication in a cloning vector.

5.1.1.2 MULTIPLE CLONING SITE (MCS)/POLYLINKER SEQUENCE

All cloning vectors must have a unique site where the exotic DNA strand (gene of interest) can be ligated or removed, which is termed as multiple

cloning sites (MCS). An MCS, also known as polylinker sequence, carries various (approx. 20) restriction sites which is recognized by various restriction enzymes. These enzymes are used to introduce a space for the desired gene to ligate within the vector as well as to remove a target gene from a vector or any other construct (Carter and Shieh, 2015).

MCS site is located downstream to the promoter region and it is engineered in such a way that a single restriction site is only once available in the whole vector sequence. This allows the restriction enzymes to only cut at a single site of the vector, and the corresponding enzyme would be called single cutter for that specific vector. A hypothetical MCS site which is recognized by various restriction endonucleases (ER) within a cloning vector is shown in Figure 5.2. You can see a list of restriction enzymes, which are representing the existence of their restriction sites at the MCS region in cloning vector. Now if we restrict this cloning vector with *NdeI* and *NcoI* enzymes, the vector will be cut only at these two sites, and we must cut the gene of interest with the same two enzymes in order to insert a target gene with in this region. We will later discuss restriction enzymes in this chapter, in detail.

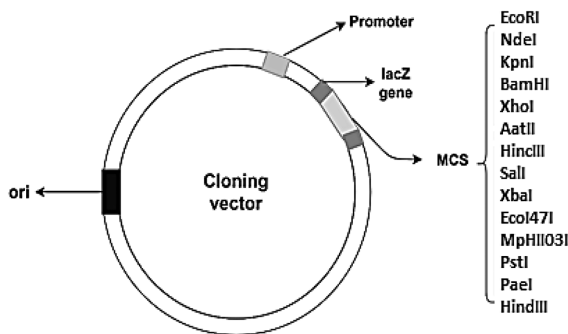


FIGURE 5.2 Multiple cloning site in a cloning vector.

Scientists engineered these sites to achieve their desired results, e.g., pUC18 is an engineered cloning vector with a unique polylinker sequence, scientists have reversed its polylinker sequence and named the resultant vector as pUC19 and the introduction of two more restriction sites, i.e., *NcoI* and *NdeI* in them give rise to a new cloning vector, pUC72 (Xavier et al., 2009). These changes are made to impart useful characteristics to the cloning in order to make them diverse in their functionality.

5.1.1.3 SELECTABLE MARKER GENE

A selectable marker gene is another necessity of a cloning vector as it helps in identification of the transformed cells among the population which carry the vector. For instance, an antibiotic resistance gene is usually present in the cloning vector as a selectable marker, and only those cells that carry the vector can grow in the presence of a respective antibiotic. There are lots of antibiotics and different vectors carry resistance genes against one or more antibiotics. A vector with a specific antibiotic resistance gene (e.g., ampicillin) can only grow in the presence of that specific antibiotic (ampicillin) and other untransformed cells would not grow in the presence of the antibiotic, which lays the foundation of selection of successfully transformed cells. However, the presence of any other antibiotic (e.g., kanamycin, chloramphenicol, etc.), besides ampicillin will kill all the transformed and untransformed cells (Carter and Shieh, 2015).

5.1.1.3.1 How It Works?

Here we take ampicillin resistance (Amp^r) gene as an example for understanding the function of selectable marker gene. Amp^r is a commonly used antibiotic resistance gene available in cloning vectors which catalyzes the hydrolysis of penicillin.

Ampicillin is an antibiotic, which efficiently binds to several bacterial enzymes which are involved in the formation of bacterial cell wall, hence retards bacterial growth. Molecular formula of ampicillin is shown in Figure 5.3. Amp^r gene encodes for beta-lactamase, an enzyme which hydrolyzes the beta-lactam ring of ampicillin, hence, detoxifies the drug. So, if we grow bacterial cells in a culture with ampicillin in it, only those cells will grow that have amp^r gene in them, which in other terms are successfully transformed cells.

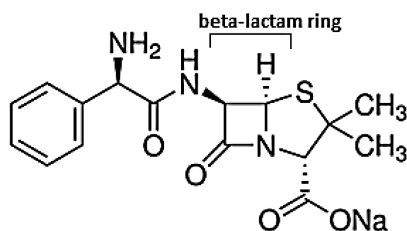


FIGURE 5.3 Molecular structure of ampicillin.

5.1.1.4 REPORTER GENE

In some cloning vectors, a reporter gene is also incorporated which helps in further screening of desired transformants (cells which have undergone successful transformation with desired gene). LacZ gene is widely used reporter gene in cloning vectors which encodes an enzyme, i.e., beta-galactosidase which hydrolyzes lactose into glucose and galactose, and in blue-white screening (Suzuki & Griffiths, 1976). This property of enzyme has been exploited by the scientist to make it a reporter gene in cloning vectors. Naturally, lactose acts as both; the inducer of the lac operon, as well as the substrate for the lacZ gene. However, in blue white screening, X-Gal is used along with isopropyl β -d-1-thiogalactopyranoside (IPTG) to carry out the selection process. IPTG is a synthetic analog of allolactose, which acts as an inducer instead of lactose for ensuring the expression of the lacZ gene. Similar to lactose, IPTG works as a substrate for beta-galactosidase, but unlike lactose, it is not metabolized into glucose and galactose. X-Gal, on other hand, is metabolized by beta-galactosidase to produce galactose and 5-bromo-4-chloro-indoxyl. The latter one gives blue color upon dimerization (Brandt et al., 1991).

Reporter gene (i.e., LacZ gene) occupies the MCS of a vector as shown in Figure 5.4. If the desired gene has not been inserted at the MCS site of the cloning vector, it means that the reporter gene is intact and able to

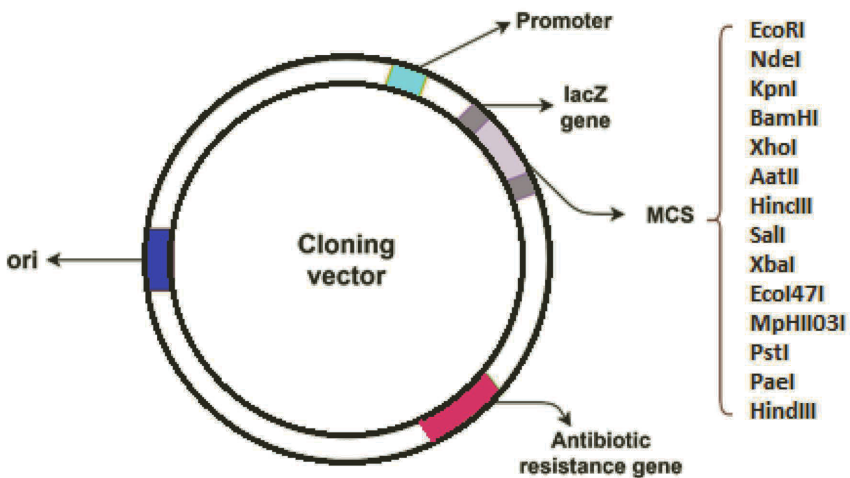


FIGURE 5.4 A functional cloning vector with Ori, MCS, selectable marker, and reporter gene.

perform its function, i.e., the production of blue colonies. On the contrary, if the gene of insert has been successfully inserted at the MCS site, it will disrupt the beta-galactosidase gene and white colonies will appear after transformation. So, if a cloning vector without our gene of interest has been transformed in the cell, it will grow on in the presence of ampicillin but produce a blue colony, whereas, if a cell has successfully transformed with a vector carrying our gene of interest, it will grow in the presence of ampicillin and produce white colony.

5.1.1.4.1 How Blue Colonies Form?

In blue-white screening, cells are grown in culture with X-Gal in it, which is hydrolyzed by the enzyme (beta-galactosidase) into galactose and 4-chloro-3-bromo-indigo (gives blue color upon dimerization), as mentioned earlier. Cells with intact lacZ gene (cells without gene of interest) can only produce blue colonies, whereas, if lacZ gene is somehow attenuated (by the insertion of desired gene), it would fail to produce beta-galactosidase and subsequently white colonies would appear as illustrated in Figure 5.5.

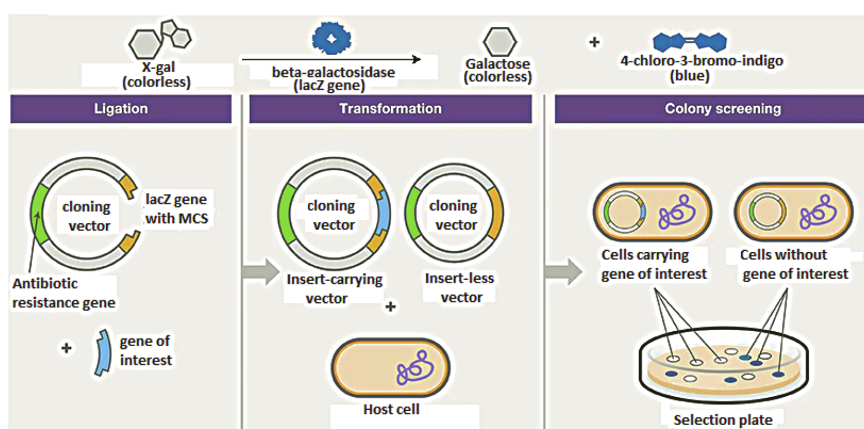


FIGURE 5.5 Principle of blue-white screening.

5.2 TYPES OF CLONING VECTORS

There are various types of cloning vectors with respect to the size of desired gene which is to be inserted in cloning vector. A brief list of such cloning vectors is given in subsections.

5.2.1 PLASMIDS

Plasmid is a small, autonomously-replicating extra-cellular DNA molecule which is present in the cell, besides chromosomal DNA. They are mostly found in bacteria whereas, some archaea as well as eukaryotes are also known to have plasmids. These are the most commonly used cloning vectors for the synthesis of recombinant cells. Plasmids have the capacity to efficiently clone approx. 10 kb of gene of interest. Any gene below this size can easily be cloned and manipulated using this system however, plasmids are not efficient to clone larger size DNA inserts (Lodish & Zipursky, 2001). Bacteria are often used as host cells for plasmids; however, cell lines can also be transformed using plasmids as vectors. Copy number is a term referred to the number of copies of plasmids per cell. A plasmid copy number is directly related to the structural characteristics of its Ori point. Other contributing factors include, plasmid size and culture condition for cell growth. All these factors can be manipulated in order to design a plasmid with desired properties. pUC 19 is a genetically engineered (GE) cloning vector with high copy number (Figure 5.6). Its Ori point is manipulated in such a way as to produce 500–700 copies of plasmids per cell.

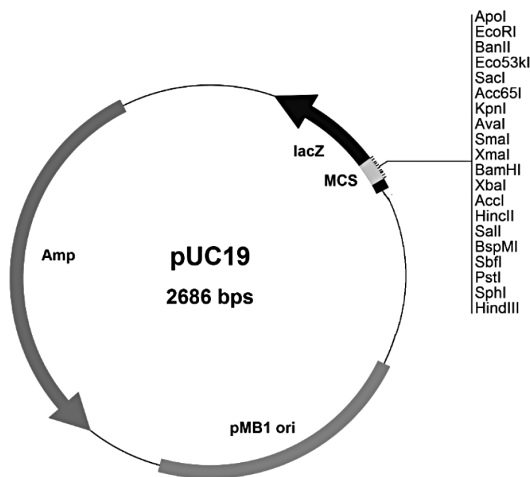


FIGURE 5.6 Illustration of pUC19 vector.

Source: Adapted from Boca Scientific Inc.

5.2.2 BACTERIOPHAGE

Bacteriophage is a virus that eats on bacteria which are have various applications in rDNA technology. M13 phage and λ phage are the most frequently used bacteriophages in cloning. Phages usually contain a linear DNA molecule and unlike plasmids, single cut in DNA strand will generate two fragments. These fragments can subsequently be fused with foreign DNA elements to generate a chimeric phage. For synthesizing these chimera, foreign DNA can be fetched either from another phage or a plasmid donor carrying the gene of interest as depicted in Figure 5.7 (Chen et al., 2019).

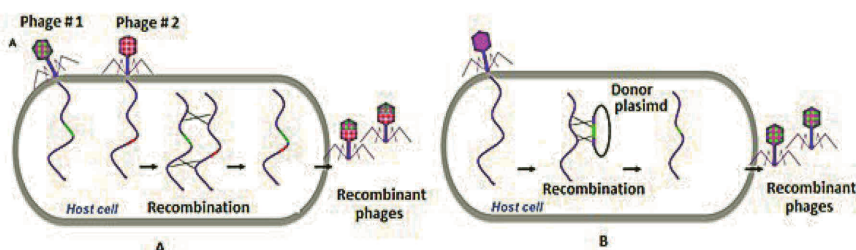


FIGURE 5.7 Image depicts how bacteriophage is used to make recombinant DNA.

Source: Reprinted from Chen et al. (2019). <https://creativecommons.org/licenses/by/4.0/>

Although it provides a reliable platform to produce chimeric constructs, the capacity of the phage head is a serious limitation as it cannot occupy larger DNA fragments. Scientists have excised the non-essential elements from phage genome to overcome this problem. For utilizing bacteriophage as a cloning vector, it should follow a lytic cycle therefore all non-essential genes would be excised from the virus, and the space created in this way could be used to insert larger DNA fragments. However, smaller size DNA molecules also failed to package properly inside the bacteriophage. This property has been used as a selection marker, i.e., if a DNA insert is too large, then only those vectors that carry the gene of interest would undergo successful transformation and vice versa. Phage vectors allow to clone approx. 20–25 kb foreign DNA fragments. Moreover, they are easy to handle and are efficiently being used in creating genomic and cDNA libraries (Casali & Preston, 2003).

5.2.3 COSMIDS

Cosmids are hybrid of plasmids and phages (Figure 5.8). They share similarity with plasmids for having an Ori, cloning site and antibiotic resistance gene. As for their likeness with phages, they encode cos sequences which allow *in-vitro* assembling of DNA into phage capsid and subsequently ease the purification process. Cos-sites are around 12 bp long, linear, single stranded DNA molecule, which are complementary to each other and creates the overhang in phages. Upon introducing them into bacteria, these cos-sites find each other form a circularize. These circulated molecules are then replicated inside bacterial cell and DNA polymerase keeps replicating the DNA as it is in circular form. This led to the development of a concatemer, where various copies of a single phage genome strung together. Host enzymes then cleave the genome at the cos site, and a single phage genome with single stranded overhangs are packaged into phage particles.

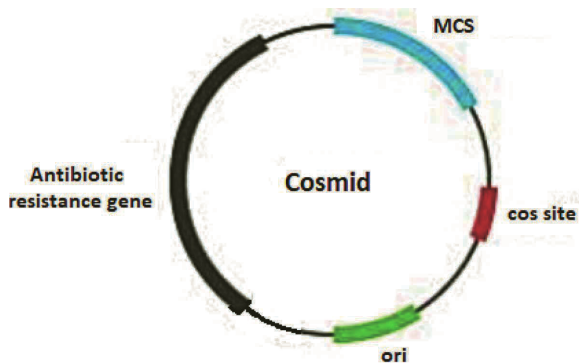


FIGURE 5.8 Illustration of a cosmid vector.

Source: Reprinted from <https://www.slideshare.net/SwatiPawar27/cosmid>

Cosmids are plasmids with cos sites in them, which contains the target gene and replicate in bacterial cells. Moreover, it lacks machinery for lytic development. Like phages, a single copy of cosmid replicates to generate a great magnitude of gene of interest which are packaged after the screening process. Unlike phages which carry 5–6 kb of phage genome, cosmids can accommodate approx. 45 kb of DNA and are widely used in generating genomic DNA libraries (Cornel, 2007). Besides its advantages, standard biological protocols are not valid when dealing with cosmids, as it is tricky

to deal with and can only be used by experienced personnel. Moreover, as the cloned fragment is larger in size when using cosmids, the risk of recombination with in the cloned sequence is also higher.

5.2.4 BACTERIAL ARTIFICIAL CHROMOSOMES (BACS)

Bacterial artificial chromosomes (BAC) are an engineered DNA molecule which is used to clone larger DNA sequences, i.e., 100–300 kb long, in bacterial cells (Figure 5.9). For this purpose, a large piece of DNA is engineered in a fashion that it circularizes itself upon ligation with bacterial plasmid DNA, followed by its replication in a bacterial cell besides bacterial own chromosome. With sufficient space and other powerful attributes of BAC, it is feasible to clone a whole gene with its regulatory element that is why BAC is extremely helpful for sequencing and mapping mammalian genomes.

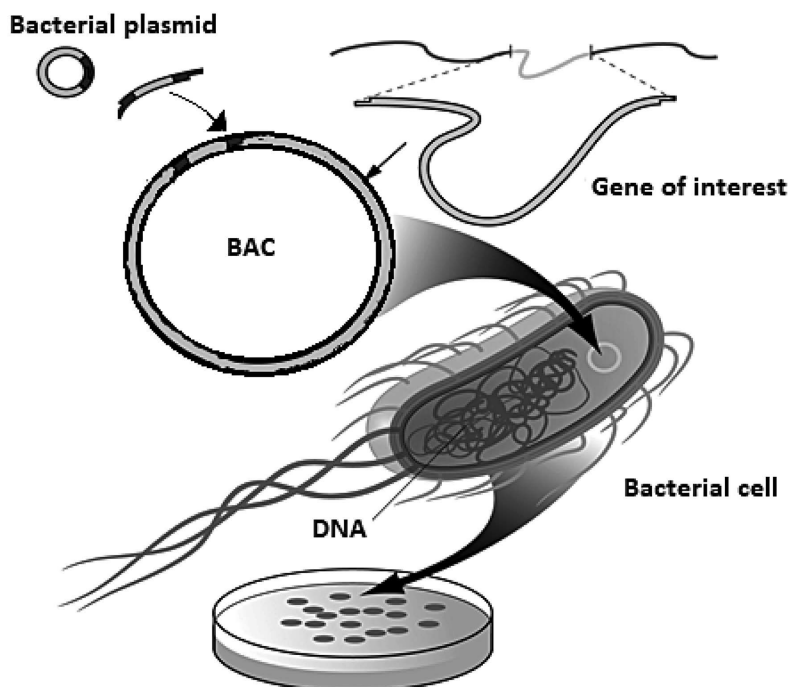


FIGURE 5.9 A schematic diagram of utilizing BAC in cloning larger DNA fragments.

Source: Reprinted from National Human Genome Research Institute.

There is a vast majority of microorganism in the biosphere and only a small fraction of them is under study because most of them are unculturable which deprive scientists from a lot of precious genome and biological pathways which might lead to marvelous discoveries. BACs provide a platform to study such unculturable microflora, their metagenomes can be indirectly studied by cloning them into BAC systems. In this way, BACs are an invaluable tool for discovering new enzymes by cataloging new genomes (She, 2003).

5.2.5 YEAST ARTIFICIAL CHROMOSOMES (YACS)

Yeast artificial chromosome (YAC) has the largest cloning capacity any known cloning vector. They are derived from the DNA of yeast, which consists of a yeast autonomously replicating sequence (ARS), a yeast telomere, two copies of yeast telomere and a selectable marker gene (Figure 5.10). The average size of DNA fragments that can be clone using YAC is 200–500 kb, whereas the maximum DNA length can be as much as 1,000 kb. Therefore, YACs are widely used to analyze the entire genome.

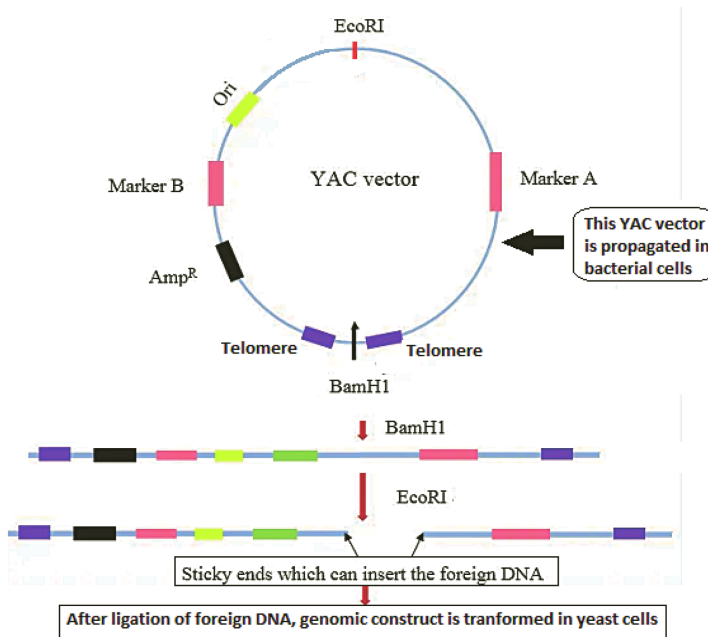


FIGURE 5.10 YAC as a gene cloning system.

Source: Partially adapted from: Orbit Biotech (2018).

YAC is initially propagated in bacterial cells as a plasmid by using plasmid Ori sequence followed by the cleavage of plasmid using restriction enzymes in between two telomere units, it will generate a single linear DNA molecule. Upon second restriction with another restriction, two DNA fragments will produce each having a selectable marker gene. At this step, the foreign DNA is inserted into these fragments followed by their transformation in yeast cells (Figure 5.10). Successful recombinants will be identified by the help of selection marker (Karcher, 1995).

YACs not only provide the large insert capacity, being a eukaryotic vector, sequences that are not stable and underrepresented in prokaryotic systems can efficiently clone in YACs. Moreover, unlike BAC and other cloning vectors YACs can be used as expression systems, and they have the tendency to perform post-translational modifications (PTMs) in eukaryotic proteins. There are a few limitations in using YACs as well. This includes high mutation rates in clones, less stability as compared to BACs, Cloning is complicated and cloning efficiency is also poor.

5.3 RESTRICTION ENDONUCLEASE

For linking of desired gene in cloning vector, the second most essential necessity is restriction ER. Restriction ER are enzymes which scan the nucleotide sequence and cut at specific sites. The sequence on the target DNA molecule where restriction enzymes cut is called restriction site. For linking a desired gene with cloning vector, we have to refine their ends such that both; gene and vector, have affinity for each other. Take it as if someone asks you to make a jigsaw puzzle, you have to refine the ends in a way that user joins the corresponding ends in order to solve the puzzle. In molecular cloning, this purpose of refining ends is served by restriction ER, also known as molecular scissors, to cut/refine the ends of given nucleotide sequences.

Restriction ER are discovered in and isolated from bacteria. Approx. 4,000 of them have been identified so far, which recognize more than 250 different DNA sequences (Roberts et al., 2015). In bacteria, these enzymes cut the foreign DNA of infectious sources, i.e., bacteriophages, and thus involve in bacterial defense mechanism. These enzymes do not recognize or cut bacterial own DNA, as the restriction sites in bacteria are modified. This modification mainly involves methylation of one or more nucleotide

bases at the recognition site, due to which these sites no longer act as substrate for these enzymes, keeping bacterial genome safe.

Restriction ER cut a single DNA strand at a time. As DNA is a double stranded molecule, to make a precise cut on both strands, restriction enzymes mostly recognize palindromic sequences as their recognition site and cleave within or very close to this site. A palindromic sequence is a nucleotide sequence in which the arrangement of nucleotide is identical in both strands, when reading both strands from the same direction, i.e., 5'-3' direction, e.g.:

5' AGCTGGCCTTGGCCAGCT 3'

3' TCGACCGGAACCGGTCGA 5'

This sequence is a palindromic sequence and you can see that nucleotide sequences on each strand is identical in 5'-3' direction as well as the enzyme recognition site. In this way, the palindromic sequence provides identical restriction sites to a restriction enzyme on both strands of DNA, in order to make a precise cut at the target site.

Upon restriction, the restricted site, a 5'-phosphate group and a 3'-hydroxyl group on each stand is formed. As DNA is a double-stranded molecule, restriction leads to the production of free ends, which have the capacity to self-ligate or ligate with any other DNA molecule that has corresponding flanking ends. Two different DNA molecules when cut with the same restriction enzyme have the ability to ligate with each other (in the presence of DNA ligase) to form rDNA.

5.3.1 TYPES OF RESTRICTION ENDONUCLEASE

There are four types of restriction ER which are classified on the basis of their cleavage sites and required cofactor (Table 5.1). Among all restriction ER, type II restriction ER are widely used in rDNA technology, because they cleave close to or within the recognition site, hence it allows the scientists to produce specific ends within a short stretch of DNA. Moreover, unlike type I and type III restriction ER, they do not require any ATP to carry out their function. Although type IV restriction ER also have similar characteristics but it is used to target modified DNA molecules such as methylated, hydroxymethylated, and glucosyl-hydroxymethylated DNA. A brief detail of each type of restriction ER with a number of subunits they carry, cofactors they require and their cleavage sites are listed in Table 5.1.

TABLE 5.1 Types of Restriction Endonucleases

Types	Subunits	Co-Factors	Cleavage Site
Type I	Multi-subunit	ATP, AdoMet, Mg ²⁺	Cleavage sites are far away (approx. 1,000 bp) from recognition site
Type II	Single subunit	Mg ²⁺	Cleavage sites are within or at short specific distances from the recognition site
Type III	Multi-subunit	ATP, Mg ²⁺	Cleavage occurs at sites 25–27 bp from the recognition site
Type IV	Single subunit	Mg ²⁺	Cleavage site is close to or within the recognition sequence

5.4 TYPES OF ENDS PRODUCED AFTER RESTRICTION DIGESTION

On the basis of the flanking ends of DNA molecule produced after restriction, ER can be classified into two types, i.e.: blunt ends and sicky ends generating restriction enzymes (Leonard, 1986).

5.4.1 BLUNT ENDS

Upon digesting a DNA fragment with a restriction enzyme, if both ends of DNA fragments terminate in a base pair, then such an assembly of ends produced are called blunt ends. They are universally compatible with other blunt-ended DNA. Blunt ends are generated by various type II restriction ER, e.g., *EcoR* V.



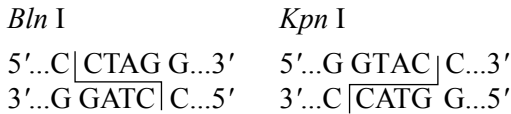
In this example, GATATC is the recognition site for *EcoR* V whereas, the red line is representing the cleavage point for *EcoR* V.

A serious disadvantage of using blunt end-producing restriction enzymes is the misorientation of the desired gene in the cloning vector.

5.4.2 STICKY ENDS

Some restriction ER introduces a staggered end and produces sticky ends/overhangs. The overhangs are unpaired, single stranded stretch of DNA nucleotides that have affinity with corresponding nucleotides. Unlike blunt ends, misorientation of gene of interest is not a problem with these ends. The sticky ends possess 3' and 5' overhangs of 1–4 nucleotides,

these small overhangs are called cohesive ends. 5' and 3' cohesive end generated by *Bln* I and *Kpn* I are shown below.



In this example, CCTAGG is the recognition site for *Bln*I whereas, on cleavage it produces C and CTAGG overhangs at 5' and 3' ends on both strands of DNA, respectively. Now this strand can only be annealed with a DNA strand which carry CTAGG as its 3' flanking end and C as its 5' overhang in the presence of DNA ligase, which means it has the potential for self-ligation. Similarly, another restriction enzyme, i.e., *Kpn*I is also shown which have a different restriction site and generate C'-overhangs on both strands.

For introducing a gene within a cloning vector, if we cut that vector and gene with a single restriction enzyme, it would be termed as single digestion. In this type of digestion, a significant fraction of vector and gene would be wasted, due to self-ligation of vector with vector, and only a small fraction of vector and gene would be ligated to form desired construct. To overcome this problem, scientists use a double digestion method where, they cut vector with two different restriction ER, similarly the gene of interest is also digested with the same two enzymes. In this way, both vector and gene lose their ability of self-ligation because now both ends of a vector as well as the gene lack complementarity with each other and only a vector and gene with complementary ends have the ability to ligate with each other. Another problem while using a single restriction enzyme is the misorientation of the ligated gene, which are shown by reverse directionality of arrows in Figure 5.11. Although this problem is more serious in blunt end producing enzymes and is significantly reduced by using sticky end producing enzymes, but chances of misorientation are still present with the latter one. Double digestion solves this problem and produces a vector with our gene of interest in a perfect orientation (Griffiths et al., 2000).

To recapitulate, double digestion has several benefits over single digestion. In case of single digestion, we get three types of plasmids, a self-ligated vector, a vector with gene of interest in reverse orientation and a vector with gene of interest in proper orientation. Among these only the last one is the desired one and the other two indicate the waste of time and resources. On the other hand, double digestion only produces a single type of plasmid, i.e., vector with the gene of interest in a proper orientation, which is extremely helpful in downstream selection processes.

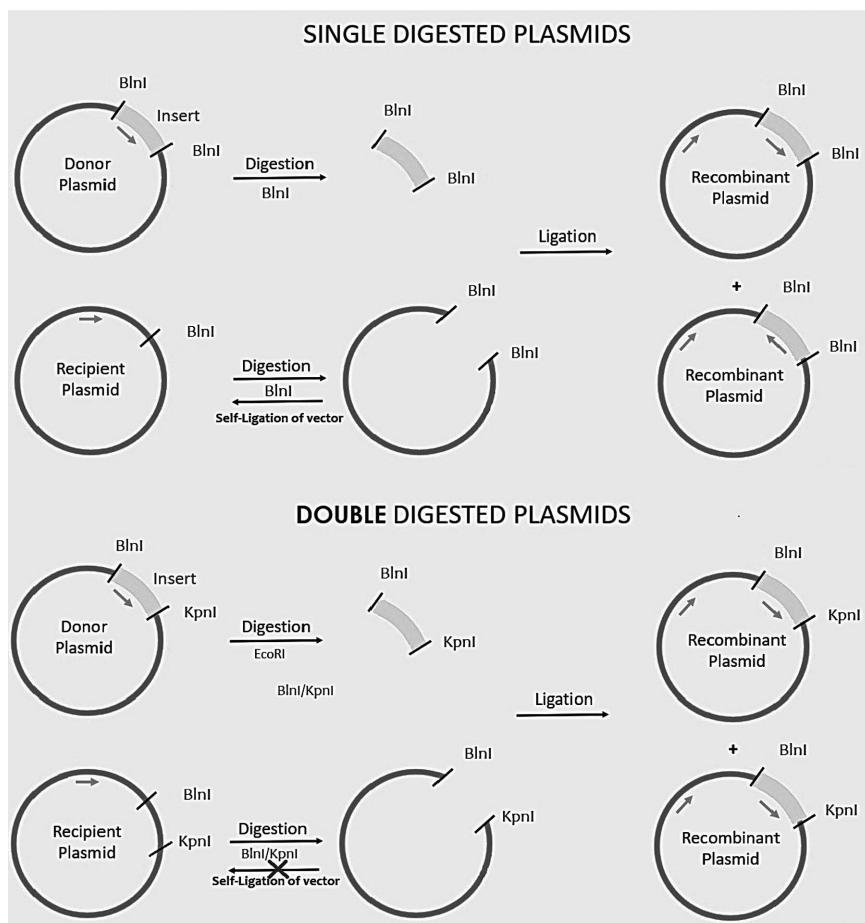


FIGURE 5.11 Single digestion versus double digestion.

5.5 DNA LIGASE

The DNA fragments formed by the application of restriction ER are ligated together using another enzyme called DNA ligase (Figure 5.12). DNA ligase carries out the ligation process by catalyzing the formation of a phosphodiester bond between 5'-phosphate group of one strand with 3'-hydroxyl group of the other strand. DNA ligase has a vital role in DNA repair and DNA replication processes in living organisms. In rDNA technology, DNA ligase is used in gene cloning to ligate gene of interest in cloning vectors to form rDNA.

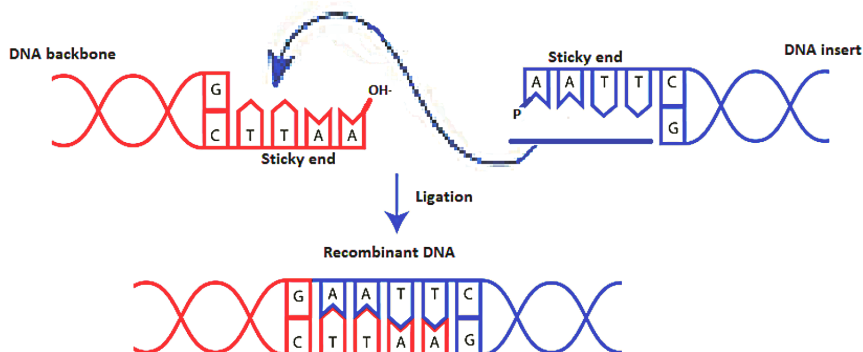


FIGURE 5.12 An illustration of how DNA ligase ligate two DNA fragments.

Source: Reprinted from Addgene. <https://www.addgene.org/protocols/dna-ligation/>

T4 DNA ligase, isolated from bacteriophage T4, is widely used in rDNA technology. It can efficiently ligate both sticky end and blunt-end DNA fragments however, it can only catalyze a ligation reaction in a double stranded DNA, and unable to do so for a single stranded DNA molecule. Moreover, it uses ATP as a co-factor and performs optimally at 16°C. Scientists have developed various mutants of T4 DNA ligase to improve its *in-vitro* activity and to make it more sensitive to UV radiation and alkylating agents.

Ligation of sticky end fragments is highly efficient due to the presence of overhangs whereas, it is less efficient in the case of blunt-end fragments. In order to improve the ligation efficiency of blunt end DNA molecules, various sequences like linkers, adaptors, and homopolymer tailing are introduced in the DNA followed by their conversion to overhangs which led to successful ligation to form a recombinant construct. The complementary sticky ends come across each other in a reaction vial, and form hydrogen bond. Later on, ligase act to form phosphodiester linkage between the deoxyribose sugar molecules.

5.6 LINKING OF GENE INTO A CLONING VECTOR

In rDNA technology, a cloning vector, gene of interest, restriction endonuclease and DNA ligase are vital necessities to insert a gene into a cloning vector. After getting a brief introduction about their function, now we move towards the main goal of this chapter, i.e., how to use these biological molecules to link a gene into a cloning vector.

To better understand this phenomenon, we take an arbitrary gene (i.e., insulin) and insert it into a hypothetical cloning vector with a few restriction sites, in order to get an insight for making a rDNA. First of all, we will find two common restriction sites in both gene and vector as we aim to perform double digestion on them, to prevent misorientation of gene in vector as well as to obtain a higher yield of successful transformants. If a gene lacks a specific restriction site, we can introduce one by performing a simple PCR with primers that contain the sequence of desired restriction site. In this way, we can introduce a restriction site at either 5' or 3' or both ends of the gene as needed. As we took insulin as our specimen gene, we make sure that both insulin gene and cloning vector have identical restriction site, i.e., *NdeI* (CATATG) at 5'-end whereas, another restriction site, i.e., *HindIII* (AAGCTT) at 3' end of both gene and vector as shown in Figure 5.13. Gene and vector are then restricted with their respective restriction ER separately. Successful restriction analysis is carried out through DNA agarose gel electrophoresis, followed by the extraction and purification of restricted fragments from agarose gel. Lastly, T4 DNA ligase is used to ligate the restricted fragments. This step is carried out by mixing vector and gene in 1:3 ratio (can be optimized w.r.t. the size of gene), followed by their incubation at 16°C. The ligated mixture is transformed into competent host cell and successful recombinants are screened through the selection plate. A detailed illustration of the linking of gene into a vector is shown in Figure 5.13.

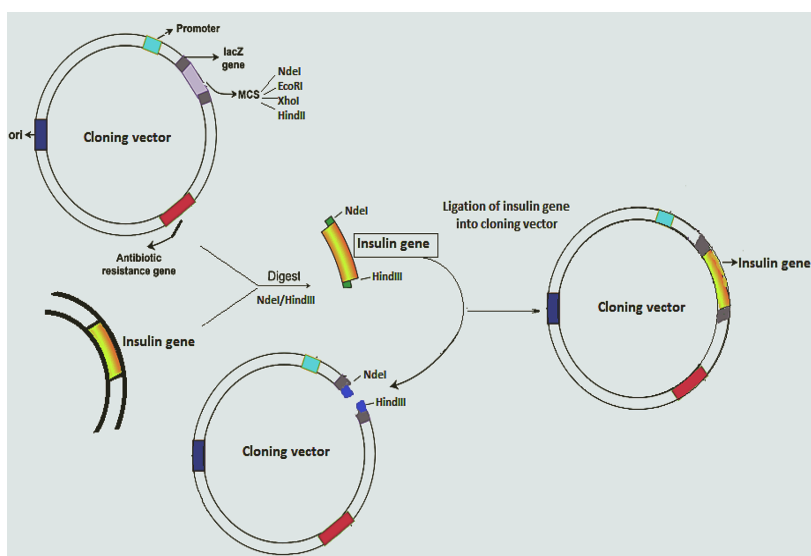


FIGURE 5.13 A systematic illustration of linking a foreign gene into a cloning vector.

KEYWORDS

- **autonomously replicating sequence**
- **bacterial artificial chromosomes**
- **isopropyl β -d-1-thiogalactopyranoside**
- **multiple cloning site**
- **origin of replication**
- **yeast artificial chromosomes**

REFERENCES

- Addgene. DNA Ligation. <https://www.addgene.org/protocols/dna-ligation/>
- Boros, I., Pósfai, G., & Venetianer, P., (1984). High-copy-number derivatives of the plasmid cloning vector pBR322. *Gene.*, 30(1–3), 257–260.
- Brandt, M. E., Gabrik, A. H., & Vickery, L. E., (1991). A vector for directional cloning and expression of polymerase chain reaction products in *Escherichia coli*. *Gene.*, 97(1), 113–117.
- Casali, N., & Preston, A., (2003). *E. coli Plasmid Vectors: Methods and Applications* (Vol. 235). Springer Science & Business Media.
- Chen, Y., Batra, H., Dong, J., Chen, C., Rao, V. B., & Tao, P., (2019). Genetic engineering of bacteriophages against infectious diseases. *Frontiers in Microbiology*, 10, 954.
- Cornel, M. E. W. B., (2007). Cloning DNA fragments. In Cornel, M. E. W. B., (ed.), *Molecular Biology and Genomics* (pp. 105–129). Academic Press.
- Del, S. G., Giraldo, R., Ruiz-Echevarría, M. J., Espinosa, M., & Díaz-Orejas, R., (1998). Replication and control of circular bacterial plasmids. *Microbiology and Molecular Biology Reviews*, 62(2), 434–464.
- Griffiths, A., Miller, J., Suzuki, D., Lewontin, R., & Gelbart, W., (2000). *An Introduction to Genetic Analysis* (p. 960). WH Freeman. New York.
- Howe, J. G., (2018). In: Nader, R. A. R. H., & Carl, T. W., (eds.), *Principles of Molecular Biology*. Elsevier.
- Karcher, S. J., (1995). Recombinant DNA cloning. In: Karcher, S. J., (ed.), *Molecular Biology* (pp. 45–134). Academic Press.
- Leonard, G., Davis, M. D. D., & James, F. B., (1986). Restriction endonucleases (REs) and their use. In: Leonard, G. D. M. D., & James, F. B., (eds.), *Basic Methods in Molecular Biology* (pp. 51–57). Elsevier.
- Lodish, H., & Zipursky, S. L., (2001). Molecular cell biology. *Biochem Mol Biol Educ.*, 29, 126–133.
- Boca Scientific Inc. *Standard Cloning Vector pUC19*. from <https://www.bocascientific.com/standard-cloning-vector-puc19-p-774.html> (accessed on 12 October 2022).

- Carter, M., and Shieh, J. (2010). Molecular cloning and recombinant DNA technology. *Guide to Research Techniques in Neuroscience* (pp. 207–227). Academic Press.
- Carter, M., and Shieh, J. (2015). Molecular cloning and recombinant DNA technology. *Guide to Research Techniques in Neuroscience* (Second ed., pp. 219–237). Academic Press.
- National Human Genome Research Institute. Bacterial Artificial Chromosome (Bac). <https://www.genome.gov/genetics-glossary/Bacterial-Artificial-Chromosome>
- Orbit Biotech (2018). YACs. *Molecular Biology*. from <https://orbitbiotech.com/yeast-artificial-chromosomes-yac-yeast-artificial-chromosomes-yac-ypc-yep-murray/> (accessed on 12 October 2022).
- Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D., (2015). REBASE—A database for DNA restriction and modification: Enzymes, genes and genomes. *Nucleic Acids Research*, 43(D1), D298, D299.
- She, K., (2003). So you want to work with giants: The BAC vector. *BioTech. J.*, 1, 69–74.
- Suzuki, D. T., & Griffiths, A. J., (1976). *An Introduction to Genetic Analysis*: WH Freeman and Company.
- Xavier, M. A. S., Kipnis, A., Torres, F. A. G., & Astofi-Filho, S., (2009). New vectors derives from pUC 18 for cloning and thermal-induced expression in *Escherichia coli*. *Brazilian Journal of Microbiology*. 40(4), 778–781.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

CHAPTER 6

Polymerase Chain Reaction

ASIMA TAYYEB and ZHUHA BASIT

*School of Biological Sciences, University of the Punjab, Lahore,
Pakistan*

ABSTRACT

Polymerase chain reaction (PCR) is a remarkable technique which can be used for amplification of DNA at exponential rate. It was invented in 1983 by Mullis. In the history of biological and medical sciences, the development of the PCR has been a turning point. PCR with its ability to amplify even a single molecule makes it a highly sensitive and qualitative technique. PCR has wide range of applications in the fields of molecular genetics, systematics, epidemiology, archaeology, anthropology, evolutionary genetics, and other branches of forensic sciences. Moreover, DNA amplified through PCR can be employed in number of processes like cloning, sequencing, sizing and analytical detection (Arnheim & Erlich, 1992).

6.1 HISTORY

The credit of invention of remarkable PCR in method 1983 goes to Kary Mullis, who was later on awarded with Nobel Prize in Chemistry for his invention in 1985 which he shared with Michael Smith. Techniques which were quite similar to PCR were in use in past, but a major breakthrough

in this field occurred with the discovery of Taq polymerase in 1976, which opened new avenues and provided opportunity for utilization of thermocycler which could be programmed (Mullis and Faloona, 1987). Taq polymerase has been derived from bacterium *Thermus aquaticus*, which are thermophilic and live at extreme hot temperature. One of main characteristic of this enzyme is that it remains active even beyond the denaturation temperature which makes it much effective and ends up the need to add enzyme at every new cycle (Chamberlain et al., 1994).

6.2 PRINCIPLE OF PCR

PCR makes it possible by *in vitro* replication, to obtain multiple copies of fragment of DNA. The fragment can be either genomic DNA, mitochondrial DNA, complementary DNA (cDNA) obtained by reverse transcription of RNA. This technique basically involves replication and it is bifurcated into three stages; denaturation stage, annealing stage and at last there is elongation stage. In PCR, products are amplified at an exponential rate (van et al., 2008).

The PCR is carried out in presence of specific reagents and it comprises of Taq polymerase, DNA template, all 04 deoxyribonucleoside triphosphate dNTPs, buffer, and $MgCl_2$ which provides co-factor for enzyme. Then tubes containing reaction mixtures are placed in a thermocycler which is programmed for successive repetition of all steps and variation of temperature ranges by Peltier effect from 0°C to 100°C. The PCR process is bifurcated as in subsections.

6.2.1 DENATURATION

As the name depicts, it involves the separation of two strands of DNA which is achieved by raising temperature. The process of denaturation is usually carried out at 95°C and at this temperature both strands separate from each other because hydrogen bonding which holds strands together is disrupted at temperature higher than 80°C and each separated strand serves as template (Breslauer et al., 1986).

6.2.2 HYBRIDIZATION

The second step in PCR is hybridization and it is carried out at temperature between 40°C–60°C and it depends upon melting temperature of primers. It is temperature where hybridization occurs between primer and template and primer hybridize with complementary region of matrix DNA. At higher temperature, more specific binding occurs between primer and template.

6.2.3 ELONGATION

The third stage in PCR is elongation which is carried out at 72°C and it involves synthesis of complementary strand at 72°C as at this temperature Taq binds with single stranded primed DNA to catalyzes replication. So, those regions of template DNA which are at downstream of primers are only selectively synthesized. In every cycle, the fragments of previous cycle act as template and right after few cycles, the most dominant species would be region where the hybridization of primers took place. After 20–40 cycles, a considerable amount of DNA is synthesized as in each cycle, the amount of DNA present in the previous cycle is doubled. The addition of final cycle is highly recommended, especially when the sequence of interest is larger than 1 kilobase (PCR Optimization: Reaction Conditions and Components, 2017) (Figure 6.1).

6.3 TECHNICAL FEATURES OF PCR

The greatest advantage of PCR lies in its speed and simplicity with which reaction can be performed. PCR reaction can be easily set up in relatively short time and multiple samples can be easily handled (Ballesteros, 2006; Bartlett & Stirling, 2003). All components of reaction mixture are added up in single microcentrifuge tube (0.2–0.5 ml) and their walls are quite thin in order to facilitate rapid transfer of heat or in some cases plates having multiple wells are employed. Moreover, some systems which are available commercially utilize capillary tubes. Once the process of amplification is completed, the product is analyzed in various ways, but real time (RT) PCR has also ended up opening the tube once the reaction has been done.

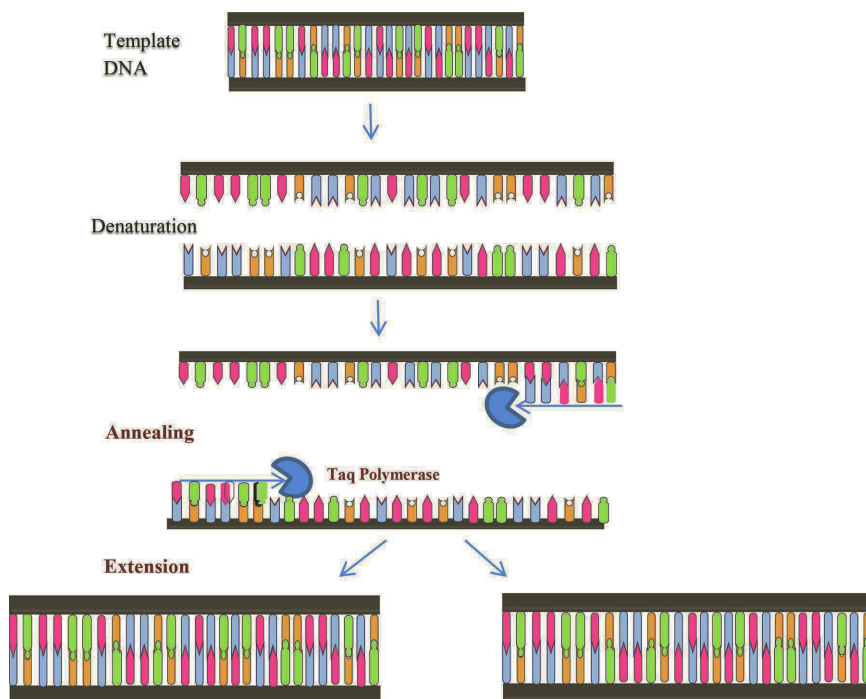


FIGURE 6.1 Bifurcation of PCR step: First, DNA template is denatured at 95°C, which is followed by annealing with primer and Taq polymerase adds nucleotides and then final extension takes place at 72°C.

6.4 TEMPLATE

One of the hallmarks of PCR is that virtually any kind of DNA can be used as a template like it may be cosmid, plasmid, lambda phage, phagemids, genomic DNA, M13 phage, or can be any other form as well. It can also be performed directly on plaques or colonies. No doubt, PCR is sensitive enough to amplify even a single copy of DNA, but in usually reaction 10^5 – 10^6 copies of DNA are added as template which is approximately equal to 1 ng of bacterial DNA, 10 ng of DNA of yeast and about 1 μ g of DNA from eukaryotic sources. The concentration of template, however is not of much concern specially in case of PCR and it works with various ranges of template concentration. While dealing with biological samples, it is sometimes necessary to remove inhibitors such as mucopolysaccharides and heme, as they can interfere with DNA.

6.5 BIOCHEMICAL COMPONENTS

6.5.1 OLIGONUCLEOTIDE PRIMERS

Primers are generally 18–30 nucleotides long and they should have 40–60% GC content. They can be used at concentration ranging from 0.1 to 1 $\mu\text{mol l}^{-1}$ with most commonly used concentration is 1 $\mu\text{mol l}^{-1}$. The very high concentration of primers promotes non-specific binding which result in non-specific amplification and consequently formation of non-specific products. T_m is defined as the temperature at which half of DNA is double-stranded, and a large number of software are available for calculation of T_m with the most commonly employed formula for this purpose is $2 \times (A + T) + 4 \times (G + C)$. However, the most careful consideration takes into account the effect of surrounding bases and ionic strength. Ideally the T_m values of two primers should be same, however the maximum difference which is allowed is of $\sim 5^\circ\text{C}$ with annealing temperature should be 5°C less as compared to T_m . No doubt, as annealing temperature approaches the T_m , amplification becomes more and more specific but yield decreases. So, irrespective of all calculations, the testing of annealing temperature is of key importance for optimization (Primer Design Tips and Tools, 2015). At the 3' end of primers, complementarity should be avoided as it promotes the formation of primer dimers (Chou et al., 1992). Such primer dimers themselves can act as template for amplification, and they can compete with actual template for polymerase, primers, and deoxyribonucleotide triphosphates (dNTPs). If it happens, then there is not only poor yield of desired product, but the further analysis may be complicated due to the formation of non-specific products (Kadri, 2019).

Oligomer formation occurs when two primers anneal with each other and their extension occurs during PCR. This double stranded primer product serves as template and during successive rounds of replication, it is also amplified and it competes with actual product. The probability of oligomerization increases with higher concentration of enzymes, low annealing temperature and at higher concentration of primers. However, the most sternness step is an initial cycle, in which oligomer formation takes place and in order to avoid the oligomerization, various stringent techniques such as Hot Strat have been described in order to maintain stringency till the point when enzymes and primers are mixed together. In Hot Start, primers and enzymes are mixed together only after stringent temperature has been reached. One way of doing Hot Start is adding

manually a small volume of buffer in which enzyme is present is present to other reaction components, which are maintained in cycler at about 80°C, and after that proceeding with normal amplification of PCR. Another method involves sealing of amplification reaction by employing wax and other reaction components are then introduced through wax barrier. This goal can also be achieved enzymatically, either through use of uracil DNA glycosylase, through DNA polymerase which is modified by blocking it with some thermolabile group, short term incubation at 50°C or by utilization of dUTP. Uracil DNA glycosylase is employed at 1–2 unit/reaction and it makes sure that any oligomer or non-specific product formed during less stringent conditions will be subjected to degradation by UNG (Diefenbach et al., 1993). This method not only causes inactivation of amplicons which are of contaminating nature but also diminishes oligomers and non-specific products. The residual UNG possesses the ability to cause degradation of any amplicon which contains U residue; therefore, PCR product must be analyzed in a short time or may be stored at 4°C for long term utilization. Besides, UNG can also be removed through organic extraction.

It is necessary to consider that the 3' end of primer should be exactly complementary to that of the template as if it does not happen then there exist high chances of ineffective amplification. Internal sequences are of not much importance. With suitable annealing temperature, pair of degenerate primers can also be used, if the sequence is highly variable or partially known. They are designed by comparing similar genes from related organisms. Whereas, sequence at the 5' end of primer is not of much importance and it should be appropriate enough to ensure hybridization of primer with template. In order to introduce restriction enzymes, non-homologous sequences can also be introduced at 5' end and it is mostly done when there is need of cloning or when other sequences such as polymerase promoters are to be added (Lawyer et al., 1989).

In multiplex PCR, it is possible to amplify various genes at the same time. Since primers of various genes have varying annealing temperatures, so in this case, optimization of reaction is of key importance (Edwards and Gibbs, 1994).

In some cases, for PCR analysis, primers are labeled with various detectable groups, such as fluorescent dyes, biotins, or radioactive isotopes. When label is attached at the 5' end, they generally do not interfere with amplification.

6.5.2 DNA POLYMERASE

Among widely available DNA polymerases, most commonly Taq DNA polymerase obtained from *Thermus aquaticus* is used. Besides, other polymerases such as *Thermococcus litoralis* and *Pyrococcus furiosus* have also been employed and low error rate has been detected with them as compared to that of Taq polymerase. Normally, in 50 μ l reaction, 0.5–2.5 enzyme units are used (Holland et al., 1991).

6.5.3 COMPONENTS OF BUFFER

For polymerase to function properly, there exists a need of monovalent salts, buffer ions and divalent cations. Most commonly employed buffer contains KCl and Tris-HCl. At 72°C, the extension temperature, the pH of buffers falls to 7.2 which is the optimal pH for polymerase. Besides, sulfate containing buffer is also widely employed and it contains Tris-SO₄ and (NH₄)₂SO₄. In order to just ionic strength, there is addition of monovalent cations. For DNA polymerase to function properly there is need of divalent cation, but higher concentration of MgCl₂ promotes non-specific binding of primers and as a result non-specific products are formed. In order to prevent adsorption of polymerase to the tube surface, a small amount of reaction buffer is also added. Some other components such as dimethylsulfoxide, glycerol, formamide, and betaine can also be added to improve yield.

6.5.4 DEOXYNUCLEOTIDE TRIPHOSPHATES (dNTPS)

dNTPs must be added carefully as excessive higher concentrations of dNTPs promote the formation of nonspecific products. Sometimes, dNTPs labeled with fluorescent, biotin, or radioactive markers are also used, and they are usually used in lower concentration as compared to unlabeled dNTPs.

6.6 TROUBLESHOOTING WITH PCR

If desired amplicon is not achieved with standard conditions of PCR, then optimization is of crucial importance. Reaction's stringency can be modulated in such a way that by altering variables, specificity can be achieved. Such as if stringency of reaction is not properly adjusted, many amplicons

of varying lengths can be produced, whereas if conditions would be too stringent, then no amplicon would be produced. At many times, it may be the most frustrating endeavor, but having good knowledge of reagents used in PCR can reduce time to a great degree. Most of problems can be solved by carefully adjusting the concentration of Mg^{2+} ions and annealing temperature. One should confirm that there is no handling error before changing the conditions. First, make sure that no reagent was contaminated and also note that were their primer dimers before or any non-specific products. GC contents should also be carefully estimated (Roux, 1995).

First, determine that which reagent is in actual causing trouble and for this there is need to prepare new reagents, then add one new reagent every time in order to rule out the problem. This will demonstrate which reagent was catastrophic. When DNA becomes old, then there is a chance of accumulation of inhibitors, so in that case addition of BSA also works.

Second, check primer dimers as they form when primers are self-complement or they anneal with each other instead of annealing with each other. In this scenario, check out the concentration of primers as an excessive amount of primers can result in the formation of dimers. Utilization of Hot Start Thermo-cycler or addition of DMSO may also work. Otherwise, there is a need to design new primers.

PCR products will not form when reaction conditions are much more stringent, Hairpin loops and other primer dimers formed during reaction also prevent the formation of product because they lose their ability to anneal with the template.

GC contents must also be analyzed carefully as GC contents greater than 60% pose a serious challenge, but this trouble can be ruled out through the use of various additives (PCR). National Center for Biotechnology Information [Online]).

6.7 FIDELITY

The fidelity of DNA polymerase is of key importance in the case of Taq polymerase, especially in case of single base substitution. It governs similarity between template and PCR product. Sometimes, a wrong base is incorporated by Taq polymerase and then it is promulgated during subsequent cycles. Higher will be fidelity of enzyme, more will be resemblance of sequence of template with that of product (Wages, 2005). A wide range

of commercially available enzymes and DNA polymerases have shown considerable fidelity. Regardless of intrinsic differences among various enzymes, fidelity is affected by a wide range of extrinsic factors. Sometimes, low fidelity conditions are used to introduce random mutations (Goodman, 1995).

6.8 VARIOUS MODIFICATIONS OF PCR

6.8.1 TOUCHDOWN PCR

A PCR in which each cycle differs from other by 1°C and each cycle is repeated twice with temperature ranging from 10–20°C in 20–40 cycles and during the process, the touchdown temperature reach and pass. The basic process may be changed with respect to temperature drop, temperature range and cycle numbers. In this case, preference is given to reaction which has highest T_m because it will be more specific.

Like if some amplicon is produced at $T_m \sim 4^\circ\text{C}$ higher than other one, then preference will be given to it as temperature drops down by 1°C in every cycle, and quantitative advantage of initial product formed at higher temperature will be $2^{4 \times 2}$ (256).

This method mainly finds its way when primers T_m is unknown or when multiplexing of amplicon is to be done with primer sets by employing various annealing temperatures (Osterrieder et al., 1994).

Touchdown PCR is mostly employed when sequence of gene is not known properly and there is need to employ universal primers or when there is need to amplify genes from related species or when sequence of gene has been deduced from amino acids and perfect nucleotide sequence is not known. It also finds its application when mismatches are not a major issue and, in this case, less stringent conditions are required as compared to traditional PCR (Roux, 1994).

6.8.2 ALLELE SPECIFIC PCR

A revolutionary approach and concept initiated by Newton and his colleagues is being employed for detection of single nucleotide polymorphism (SNP)

is allele specific PCR also known as PCR amplification of specific allele or amplification refractory mutation system (ARMS) (Newton et al., 1989). In allele specific PCR, specific primers are designed which allow amplification to proceed only if 3' end of primer is perfectly complement to base at variant or wild type sequence. Once the PCR is done, analysis of amplified product on gel electrophoresis allows detecting and helps in differentiating SNPs (Kwok & Chen, 2003). To detect specific product of PCR, several robust and innovative methods have been employed such as analysis of melt curve in which there is need of specific probes which are labeled or probe hybridization in which there is need of probes labeled specifically. Allele specific PCR has been employed widely in many areas such as microbiology, genetic disorders, pharmacogenetics, and others (Lee et al., 1993).

Allele specific PCR offers a relatively simple method for detection of SNPs and optimization of PCR methodology and design of primers are crucial steps. After the protocol is optimized, the execution is as simple as of conventional PCR.

Besides, AS-PCR several other methods, such as pyrosequencing, high resolution melting, PCR-RFLP, and probe hybridization can also be used for detection of SNPs but AS-PCR offers the most advantages. In RFLP-PCR, there is need of incubation with specific restriction enzymes which allows digestion and separation of restricted fragments. Whereas other methods no doubt are much faster and robust but they are quite expensive and there is need of specific reagents and highly sophisticated instrumentation (Sapkota et al., 2008).

6.8.3 IN SITU PCR

Cellular materials present in solution or adhered to various surfaces can be subjected directly to *in situ* PCR. The protocol for *in situ* PCR takes into account the kind of specimen, fixative, and methods used for processing. In case when sample is embedded in paraffin, the paraffin is removed and sample tissue is rehydrated before initiation of reaction where in case when components are already in liquid form, they are directly added to tube, In case when specimens are on glass slide, specimen is covered with cover slip after addition of components in order to prevent evaporation (Bagasra, 2007), For amplification, specially designed thermal cycler or

conventional PCR are employed with detection is accomplished through probes or labeled nucleotides that hybridize with amplicon. Sub-optimization in localization of target and diffusion of amplicon are main drawbacks of *in situ* PCR (Komminoth et al., 1994).

6.8.4 QUANTITATIVE PCR

Real time (RT) PCR also known as quantitative PCR made dramatic revolution in molecular diagnostic and it allowed things to be detected in RT. It allows detection, amplification, and analysis at the same time, which is of key importance in many crucial cases such as in analysis of cerebrospinal fluids (CSF) for detection and quantification of herpes simplex virus (HSV) (Ronaghi et al., 1996).

In RT-PCR, amplification is achieved in RT mainly due to the incorporation of fluorescent dye in the amplicon, which is then detected by a special thermal cycle which can read wavelength of various ranges. Fluorescent signal is produced during reaction through various ways such as by intercalating dyes like SYBR green or fluorescence labeled probes or primers (Giulietti et al., 2001; Bustin, 2000; Bustin et al., 2005).

The main difference between conventional and RT-PCR is that there is no need of post-amplification analysis and electrophoretic techniques even it ends up the need to open PCR tubes. In the case of conventional PCR, there is always a need of post-amplification process, which increases the risk of possible contaminations by many folds (Hill et al., 2002). Even, the presence of these contaminants in the environment can interfere with subsequent PCR, and ultimately it results in the generation of false positive results (Ginzinger, 2002; Wilhelm & Pingoud, 2003; Schmittgen & Livak, 2008).

RT-PCR results are more reproducible as compared to conventional PCR as they are based upon threshold of amplification cycle (C_t) in log-linear phase as compared to plateau phase of later. Besides, RT-PCR is also much more specific as it allows to analyze melting curve which ensures the purity and specificity.

Most commonly the dyes which are employed in RT-PCR like EtBr and SYBR Green are not much specific and they can intercalate in any double stranded DNA even it may be primers dimer. They can be employed in semi quantitative and qualitative analysis for analysis of melting curves,

discrimination of products and for determination of melting temperature of probes and primers. No doubt, they are also employed in quantitative analysis, but more specific results are obtained with specific probes as dyes can intercalate in any double strand and then can produce fluorescence (Bernard and Wittwer, 2002).

RT-PCR which employ oligonucleotide probes are mainly based on principle of fluorescence resonance energy transfer (FRET). In FRET, the transmission of energy from donor to acceptor occurs without emission of photon when they are in close vicinity with each other. When fluorophore is acceptor, emission of photon occurs at particular wavelength whereas, if photon is not emitted by the acceptor, then dissipation of energy occurs and fluorescence generated from donor is quenched. Some chemistries of fluorophores are as in subsections.

6.8.4.1 MOLECULAR BEACON

The long probes having quencher dye at 3' end and reporter dye at 5' end. When there is no binding between probe and template, there is formation of hairpin loop structure and dyes are juxtaposed as a result the fluorescence is quenched. The probe is complementary to that of template sequence, when hybridization occurs, the reported dye is released from the quencher and fluorescence occurs (Lay & Wittwer, 1997). In case of molecular beacons, the probes and template must be more complementary to each other than internal base pairing of probe otherwise there would be formation of hair pain loops. In most of case DABCYL is employed as universal quencher with most common reporter dyes are Cy3, Cy5, ROX, carboxy-fluorescein, tetramethyl-6-carboxyrhodamine, Texas, and tetrachloro-6-carboxyfluorescein (Livak et al., 1995).

6.8.4.2 HYBRIDIZATION

In this case, two hybridization probe is used with donor dye attached to one probe and receptor acceptor dye attached to another probe. There is complementary between amplicon strands and probes, and they are designed in such a way that they position themselves adjacent to one another and they anneal internal to primers. So, in other words, there is the juxtaposition of acceptor and donor dye, hence, making the phenomenon

of FRET possible. With conventional PCR, the format of the hybridization probe is well compatible with denaturation at 95°C, followed by annealing and extension at 72°C (Gibson et al., 1996). When probes are encountered by DNA polymerase during the extension phase, their displacement rather than hydrolysis from target strand takes place which makes sure their availability for next round of amplification.

6.8.4.3 TAQMAN PROBES

These are short probes with acceptor dye attached at 3' end and reporter attached at 5' end. The acceptor dye which is attached at 3' end has function of quenching the reporter till probe is intact. During extension stage, probe is subjected to attack by *Taq polymerase* by virtue of its endonuclease activity at 5' end, hence releasing reporter from the quencher and it causes detection of fluorescence (Brandt et al., 1998; van et al., 1992). Fluorescence increases proportionally to amplicon produced during the log-linear stage of PCR reaction. Ideally such probes are designed in such a way that they anneal internally and close to primers of PCR. In order to ensure hydrolysis of the probe, the extension and annealing temperature are kept the same in case of two step PCR. As there is hydrolysis of probe, it is not available for subsequent reactions (Bassler et al., 1995).

6.8.4.4 VARIATIONS

Various variations can be introduced in RT-PCR for amplicon detection such as Lux (which employs single fluorophore attached with primer having structure of hairpin loop), Uniprimer (almost as same as molecular beacons) and Scorpion (primers which are self-probing) (Germer & Higuchi, 1999).

Some things are of key importance during optimization of RT-PCR, especially during the optimization of RT-PCR, number of PCR cycles versus fluorescent signal's growth curve must be monitored with utmost care. The curve should be sigmoidal and must be composed of three phases; first phase is also known as lag phase and in this phase, there is observation of only background signal. This is followed by exponential phase and then plateau phase.

Early cycles of amplification are represented by the baseline phase during which amplification is at initial stages and threshold has not crossed. During this stage, the fluorescent signal is not specific, and it mainly arises from autofluorescing components of reaction and by unbound probes (Ferré et al., 1994). The exponential growth occurs during log phase of reaction with C_t representing the number of cycles required to cross the threshold and point at which fluorescence crosses the background noise. Inverse relation exists between template copies and C_t value. Greater will be template, sooner the C_t will cross background, whereas, with smaller amount of template C_t value would be much higher (Syvänen, 2001).

Efficiency of amplification can be determined through measuring the slope of the log-linear phase. The slope can be estimated through identification of C_t of standards with known value and then linear regression line is plotted. $E = 10^{-1/\text{slope}}$ is used for calculation of efficiency and slope value should be in between -3 and -4 . When there is transition of log-linear curve from positive to negative, there is observance of inflection point as it enters in plateau phase and this point is of crucial importance. When there is no observance of inflection point, then it means that signal drift rather than DNA amplification was represented by curve, making it key step in interpretation of test.

The main characteristic of drift is slowly augmentation or depression in fluorescence without product amplification. All software is equipped with tools that help in further calculations and analysis.

When reaction enters into plateau phase, most of the components are already consumed, so in this step amplification either does not occur or very slowly occurs, thereby fluorescence signal is also either much low or ends up and curve levels. RT-PCR uses much sophisticated software which can quantify the product in log-linear phase thereby it makes quantification much easier and results are reproducible. In RT-PCR, the concentration of unknown samples can also be determined by generating standard curve from known standards and then this curve is used to determine concentration of unknown sample by its C_t value and making intersection. Instrumentation of RT is much versatile and diverse (Figure 6.2).

In RT-PCR, melt curve analysis of product can also be performed. In RT-PCR, the amplicon undergoes melt curve analysis with which labeled probe has been attached. T_m is defined as the temperature at which half of double stranded is converted into single stranded DNA. T_m of DNA enriched in GC is higher as compared to DNA enriched with AT base pairs

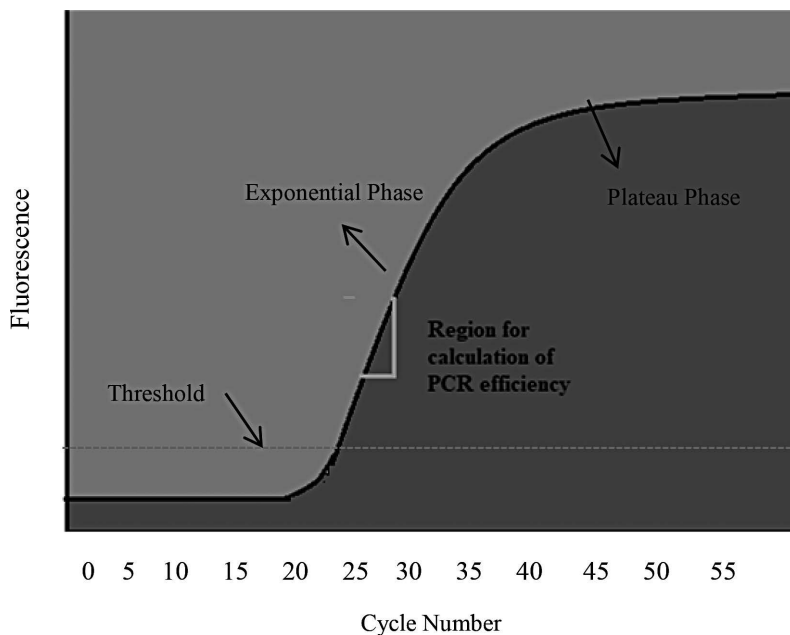


FIGURE 6.2 Interpretation of curve of qPCR: The curve plotted between number of cycles and fluorescence. Efficiency of amplification can be determined through measuring the slope of the log-linear phase. The exponential growth occurs during log phase of reaction with C_t representing the number of cycles required to cross the threshold and point at which fluorescence crosses the background noise.

because G and C are bound with three hydrogen bonds in comparison to two hydrogen bonds present between A and T. In RT instruments, the analysis of melt curve is done and after extension, there is a gradual increase in temperature in order to observe the hybrids denaturation by measuring fluorescence. Through this way, it is possible to detect even single base mutation and polymorphism as T_m of different sequences will be different.

6.8.5 REVERSE TRANSCRIPTION PCR (RT-PCR)

In a number of cases, DNA is not available template and it is RNA like in case of viral genome. In that case, RNA cannot be directly used as template and there is need to convert it in DNA which is achieved by

reverse transcriptase which is RNA dependent DNA polymerase and it causes conversion of RNA into cDNA. cDNA is much more stable as compared to RNA. In RT-PCR, the RNA which is used as starting material is degraded, and then dsDNA is produced and then amplification occurs in proceeding cycles. RNA can be isolated through kits or by using manual methods, but when this is combined with RT-PCR, then RNA analysis is done in a much better way (Freeman et al., 1999; Rajeevan et al., 2001).

It mostly finds its way in diagnosis of various viral diseases such as hepatitis C and HIV. Whereas mRNA produced by translocation like in leukemias, sarcomas, and non-Hodgkin's lymphoma can also be analyzed through RT-PCR.

6.8.5.1 RT-PCR APPLICATION; ANCHORED PCR

mRNA is used as template and it is primed with poly dT primer and first strand of cDNA is synthesized, then second strand of cDNA is synthesized by using gene-specific second primer complementary to first synthesized strand of cDNA. The main success of this method lies in its sensitivity as a single copy of RNA is also sufficient for amplification. Moreover, when the sequence of the entire gene is not known, but it is believed to contain some domain which is conserved in that family of gene, then gene amplicon can be generated successfully.

6.8.6 SOLID PHASE PCR

DNA molecules can be amplified millions of times through use of PCR, but analysis of amplicon through various methods such as Southern Blotting, Gel Electrophoresis or other hybridization techniques pose serious challenge as they all are time consuming. Another technique which offers sensitive, rapid, and specific hybridization is enzyme linked oligonucleotide sorbent assay (ELOSA) (Mercier et al., 2003). This solid phase technique of hybridization offers a highly sensitive and specific method for detection of nucleic acid in question. As the name is depicting, the technique is almost similar to the famous ELISA technique and involves the simplest steps of washing and hybridization.

6.8.7 METHYLATION SPECIFIC PCR

Gene expression in cells is regulated in many ways, such as enzyme catalyzed modification which adds methyl group at specific site. Methylation takes place at specific site also known as CpG dinucleotides, in which C is followed by G, such dinucleotides are often subjected to variety of mutations and as a result their selective depletion from genome takes place. Some of CpG regions are present in expressive genes at 5' end, and they are known as CpG islands, with most of them are found to be associated with promoters. When a CpG island is methylated, inhibition of transcription and expression of gene takes place (Lo et al., 1999).

Methylation specific PCR can be performed to address many clinical questions. First, there is treatment of DNA with sodium bisulfite as a result of which methylated cytosine is converted into uracil, hence differentiating it from un-methylated cytosine (Olek et al., 1996). Then specific primers which identify either unmethylated or methylated DNA are used, the product is analyzed on agarose gel electrophoresis and analysis for absence or presence of product in each reaction. In this way, it is possible to detect methylated, non-methylated or alleles of both kinds. Clinically, this technique has been employed for the detection of methylated CpG island in patients suffering from neoplasia, clonality presence and for analysis of genes that are imprinted.

RT-PCR is employed in the case of quantitative MSP in order to distinguish low level methylation characteristic of aging or metaplasia from high degree methylation characteristic of neoplasia patients.

6.8.8 NESTED PCR

Reactions in which there is need to increase the specificity and/or sensitivity of reaction, nested PCR is used. When there is need to amplify cDNA which has much low abundance in specimen contained of cell types of heterogeneous family or to amplify gene which is part of polymorphic gene family, nested PCR finds its applications. In nested PCR, two amplification reactions occur with two pairs of primers. As two primers are used for amplification of single locus so it is called nested PCR. In second round of amplification the product of first amplification is used as template and priming occurs by oligonucleotides which have been placed

internally to first pair of primers. The main rationale behind this strategy is that if by mistake wrong locus has been amplified, then next time the probability would be much lower due to the use of highly specific internal primers (Zhang & Ehrlich, 1994). Number of cycles can be increased to great degree through use of two pair of primers, thereby increasing the sensitivity many folds. The specificity is increased due to the fact that two pairs of primers are used to amplify the same target. It is much efficient and excellent choice for amplification of long templates, but prior knowledge of sequence is of crucial importance (Green & Sambrook, 2019).

6.8.9 PCR-BASED MOLECULAR TECHNIQUES

Microsatellites are highly variable with a number of repetitions in alleles vary to a great degree. Despite of their hypervariability, they are still employed for analysis of diversity, analysis of paternity and in order to map quantitative effect loci (Amos et al., 1999). Minisatellites share most of their features with microsatellites, with the main difference lies in a number of repetition patterns. It is anticipated that they would be replaced by more elaborative and less expensive techniques such as SNP assays (Jarne & Lagoda, 1996).

6.8.10 DNA MARKERS OF MITOCHONDRIA

Genetic diversity and phylogenetic can be analyzed through polymorphism of mitochondrial DNA (Nijman et al., 2003). DNA follows maternal pattern of inheritance and in its recombination does not take place. These specific features allowed biologists to analyze mtDNA for the construction of inter and interracial evolutionary relations by following the patterns of mutations. Specific mtDNA tags provide a quick approach to identify hybridization between species. Wild ascendants can also be traced through polymorphism in the regions which are hypervariable (Nijman et al., 2003).

6.8.11 SINGLE NUCLEOTIDE POLYMORPHISM (SNP)

To study genetic diversity, SNPs are also used as markers. SNPs carry little information as they are biallelic and large information must be employed in order to reach a panel of 30 microsatellites (Myakishev et al., 2001).

With the evolution of automation in molecular techniques, it is anticipated that in the near future, the cost to type SNPs would be reduced to a great extent. SNPs not only allow joint analysis but also direct comparison of various experiments (Syvänen, 2001). In the future SNPs would be used to study genetic diversity as they can be used easily in the analysis of neutral or functional variations (San Cristobal et al., 2006). Various experimental procedures such as DHPLC, SSCP or sequencing can be used for generation of SNPs. It is not possible to apply standard genetic parameters of population if data is obtained randomly like when initially a small population was selected for identification of SNPs and later on typing was on a large proportion of population (International Chicken Polymorphism Map Consortium, 2004; Nielsen & Signorovitch, 2003; Clark et al., 2005).

6.8.12 AMPLIFIED FRAGMENT LENGTH POLYMORPHISM (AFLP)

The most common biallelic markers are AFLP (Ajmone et al., 2002). Variations present in many loci can be simultaneously arranged and then nucleotide variation of genomic regions which are unknown can be detected (De Marchi et al., 2006), with mutation may be in functional genes which are not determined. Their mode of inheritance is dominant and it poses a serious challenge as it suppresses their ability to document genetic analysis of interracial population (Negrini et al., 2006) but they are still of crucial importance in order to identify not only race relations but also species which are related to each other (Buntjer et al., 2002).

6.8.13 RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP)

Restriction fragment length polymorphism (RFLP) employs restriction enzymes that cut at specific sites and the most commonly they are employed in downstream PCR to identify alleles which differ from each other at site of restriction (Amos et al., 1999). First, PCR is used for amplification of gene fragment and then amplified fragment is restricted through specific restriction enzyme which cut only one specific form of allele. Electrophoresis is then employed for detection and analysis. Microsatellites are quite smaller in number and size and they can be amplified easily from

DNA extracted from samples and then can be analyzed for polymorphism through sequencing (Jarne & Lagoda, 1996).

6.9 APPLICATIONS OF PCR

6.9.1 CLINICAL DIAGNOSIS

In clinical applications, the PCR finds its way for diagnosis of diseases and the simplicity, sensitivity (D'Aquila et al., 1991) and specificity of PCR has just revolutionized the field of clinical diagnosis (Mirasena et al., 2008). Bacteria which cannot be cultured on artificial media or are difficult to culture can be easily detected through PCR. *Mycoplasma genitalium* pathogenicity can be detected through PCR. At least four organisms can be detected through the employment of this method. *Treponema pallidum*, the causative agent of syphilis can be easily detected through PCR through amplification of gene responsible for coding of membrane protein. *Mycobacterium tuberculosis*, the main agent of tuberculosis affects more than one third of the population worldwide with prevalence mostly in developing countries. Recently, in developed countries, it has re-emerged, especially among the homeless and poor, and as a result, a number of strains with antibiotic resistance are emerging. In traditional methods it takes long time spanning up to several weeks to cultivate this bacterium but PCR provides rapid and quick approach for rapid diagnosis of disease (Edwards & Gibbs, 1994).

In the case of viral diseases, the amplification of viral genes forms the foundation for the detection of various kinds of viral pathogens. HIV is the main causative agent of AIDS, and it attacks the immune system. Conventionally, it is detected by immunological ways but it takes several weeks for antibodies to develop and during early-stage patients will be detected negative through conventional tests whereas virus can be easily detected through PCR. In this case, the RNA genome of the virus in plasma or proviral DNA integrated into leukocytes genome can be detected through RT-PCR. This identification has not only played a key role in diagnosis but has also enhanced the safety of blood banks by many folds. PCR techniques have also been employed for identification and epidemiological studies of other kinds of retroviruses such as human T-lymphocyte viruses and they have successfully detected human papillomaviruses, which can cause diseases of diverse nature. PCR techniques can be used to easily

distinguish more harmful strains which can cause even cervical cancer from those of harmless strains (Espy et al., 2001).

The detection of enteroviruses, which causes serious intestinal diseases in children by conventional methods pose serious challenges because of conventional methods cannot detect low level of viruses but this problem can be handled easily through use of PCR, which through its enhanced sensitivity and specificity can detect virus even at very early stage. PCR has also been found successful for detection of HSVs, human parvoviruses, adenoviruses, rotaviruses, Epstein Barr viruses, cytomegaloviruses, coronaviruses, and many others. Besides having the ability to detect the PCR can also be used to differentiate between various strains and in this way, more specific treatment is possible.

Many diseases, such as malaria and leishmaniasis, are caused by parasites, and at present, malaria poses a threat to more than 40% of the world's population with a number of resistance strains emerging day by day. Apart from diagnosing various viral diseases, PCR can also be used to distinguish between wild type and resistant strain (Innis et al., 1999).

A number of diagnostic kits based upon PCR are now commercially available. Diagnostic kits for Hepatitis C virus, *M. tuberculosis*, HIV, *Chlamydia*, and many other viruses have been introduced by Hoffman-La Roche. It is anticipated that soon diagnostic kits or detection of a large number of viruses would be commercially available. It is considered as complex test in laboratory despite of rapidity and simplicity associated with it. The expenses of commercialized kits and well-trained personnel with false positive results pose serious challenge. But automation of assay and highly sophisticated technology will make it common in laboratories for diagnosis (Innis et al., 2012).

6.9.2 PRENATAL DIAGNOSTICS

Various genetic disorders can be identified through prenatal diagnosis which is performed on amniotic fluid taken through chorionic villus sampling (CVS) or by amniocentesis. Conventional testing on these fluids takes more than one week, but PCR has revolutionized this field and result comes within a single day. If the option to abort is being taken, then time is of crucial importance and PCR is highly time effective technique. PCR made the development of non-invasive techniques for prenatal sampling

possible as risk is associated with CVS and amniocentesis so, it is preferable to move toward less invasive techniques. A number of cells from the fetus circulate in maternal blood and techniques have been demonstrated to isolate and purify them. Polymorphism and mutations on the Y chromosome can be detected by taking a maternal blood sample, and in that case, there is no need to purify fetal blood cells. It is anticipated, that in the future, a number of genetic diseases would be easily diagnosed through PCR (Markham, 1993).

6.9.3 MOLECULAR GENETICS

Besides being valuable and most effective tool for diagnosis and detection of large number of infectious diseases, PCR also finds its application in detection of various genetic disorders. A large number of genetic disorders involve mutations and DNA fragments can be amplified through PCR and later on they can be sequenced (McPherson et al., 1991). A number of techniques such as Southern blotting, single stand polymorphism analysis, probe hybridization, hybridization of allele specific oligonucleotide probe and various others can be used for the detection of mutations. This has enabled to detect a number of mutations such as in cystic fibrosis, thalassemia, sickle cell anemia and many others. PCR can also be employed to detect genetic disorders caused by expansion of trinucleotide repeats and after PCR, they are subjected to high resolution analysis on sequencing gels (Ajmone et al., 2002). Specific thermal cycle conditions and PCR reagents have been employed for commercial test used for detection of X syndrome and it possesses the ability to amplify CGG repeats up to 800 units in same reaction which also contains normal allele with 29–31 units of repeat.

Now, PCR is used commonly for HLA typing in order to evaluate compatibility before organ or bone marrow transplantation. In transfusion medicines, the role of PCR is also under investigation.

6.9.4 FORENSICS

PCR has revolutionized the field of forensics as in past conventional methods have been employed to detect similarities or differences among various samples, and they require a massive amount of DNA to carry out

processing. PCR has the ability to amplify even a very small amount of samples found at sites of investigation or at crime scenes, so, it can be used much more effectively and efficiently to identify individuals. Now, even a small stain of blood or single hair is ample for efficient DNA analysis.

6.9.5 BASIC RESEARCH

In general laboratories of molecular biology, PCR has also proved itself an indispensable tool and it has overthrown many conventional protocols and methods (Teh et al., 2007). One of the most common techniques used by molecular biologists is cloning, in which a gene of interest is inserted into the vector and later on it is introduced into bacteria, in order to screen out transformation and to confirm which bacteria contain desired gene, PCR is employed. This method has replaced filter hybridization (Kemp et al., 1989). PCR is also indispensable in nonradioactive and radioactive labeling of fragments of DNA and for analysis of mutagenesis.

PCR also finds its applications in evolutionary and phylogenetic studies. Amplification through PCR and screening of mitochondrial DNA and genes of rRNA have been employed to deduce relation among various species. PCR can be used to investigate extinct species as it can be performed quite easily on materials of extinct species available in laboratories and museums.

6.9.6 OTHER APPLICATIONS OF PCR

PCR is extensively used for analysis of food with microbial RNA and DNA can be easily detected through this technique. Different assays for *Listeria monocytogenes*, *E. coli*, *Salmonella*, and other pathogens are in routine use to ensure the safety component of food. For genetically modified (GM) foods, the presence of modification can be assessed through primers specific for gene sequence which has been inserted. In order to detect biowarfare agents like anthrax or small pox or coliform and bacteria, environmental samples can also be analyzed through PCR. PCR finds its applications all across research due to specificity and sensitivity associated with it (Griffin et al., 1994).

KEYWORDS

- **amplification refractory mutation system**
- **cerebrospinal fluids**
- **fluorescence resonance energy transfer**
- **herpes simplex virus**
- **polymerase chain reaction**
- **single nucleotide polymorphism**

REFERENCES

- Ajmone-Marsan, P., Negrini, R., Milanese, E., Bozzi, R., Nijman, I. J., Buntjer, J. B., & Lenstra, J. A., (2002). Genetic distances within and across cattle breeds as indicated by biallelic AFLP markers. *Animal Genetics*, 33(4), 280–286.
- Amos, W., Goldstein, D. B., & Schlötterer, C., (1999). *Microsatellites: Evolution and Applications*.
- Arnheim, N., & Erlich, H., (1992). Polymerase chain reaction strategy. *Annual Review of Biochemistry*, 61(1), 131–156.
- Bagasra, O., (2007). Protocols for the in situ PCR amplification and detection of mRNA and DNA sequences. *Nature Protocols*, 2(11), 2782–2795.
- Ballesteros, E., (2006). Molecular diagnostics in coagulation. In: *Molecular Diagnostics* (pp. 311–320). Humana Press.
- Bartlett, J. M., & Stirling, D., (2003). *PCR protocols* (Vol. 226, pp. 3–525). Totowa, NJ, USA: Humana Press.
- Bassler, H. A., Flood, S. J., Livak, K. J., Marmaro, J., Knorr, R., & Batt, C. A., (1995). Use of a fluorogenic probe in a PCR-based assay for the detection of *Listeria monocytogenes*. *Applied and Environmental Microbiology*, 61(10), 3724–3728.
- Bernard, P. S., & Wittwer, C. T., (2002). Real-time PCR technology for cancer diagnostics. *Clinical Chemistry*, 48(8), 1178–1185.
- Brandt, M. E., Padhye, A. A., Mayer, L. W., & Holloway, B. P., (1998). Utility of random amplified polymorphic DNA PCR and TaqMan automated detection in molecular identification of *Aspergillus fumigatus*. *Journal of Clinical Microbiology*, 36(7), 2057–2062.
- Breslauer, K. J., Frank, R., Blöcker, H., & Marky, L. A., (1986). Predicting DNA duplex stability from the base sequence. *Proceedings of the National Academy of Sciences*, 83(11), 3746–3750.
- Buntjer, J. B., Otsen, M., Nijman, I. J., Kuiper, M. T. R., & Lenstra, J. A., (2002). Phylogeny of bovine species based on AFLP fingerprinting. *Heredity*, 88(1), 46–51.

- Bustin, S. A., (2000). Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology*, 25(2), 169–193.
- Bustin, S. A., Benes, V., Nolan, T., & Pfaffl, M. W., (2005). Quantitative real-time RT-PCR—a perspective. *Journal of Molecular Endocrinology*, 34(3), 597–601.
- Chamberlain, J. S., Chamberlain, J. R., Mullis, K., Ferré, F., & Gibbs, R., (1994). *The Polymerase Chain Reaction*.
- Chen, Q., Lu, P., Jones, A. V., Cross, N. C., Silver, R. T., & Wang, Y. L., (2007). Amplification refractory mutation system, a highly sensitive and simple polymerase chain reaction assay, for the detection of JAK2 V617F mutation in chronic myeloproliferative disorders. *The Journal of Molecular Diagnostics*, 9(2), 272–276.
- Chou, Q., Russell, M., Birch, D. E., Raymond, J., & Bloch, W., (1992). Prevention of pre-PCR mispriming and primer dimerization improves low-copy-number amplifications. *Nucleic Acids Research*, 20(7), 1717–1723.
- Clark, A. G., Hubisz, M. J., Bustamante, C. D., Williamson, S. H., & Nielsen, R., (2005). Ascertainment bias in studies of human genome-wide polymorphism. *Genome Research*, 15(11), 1496–1502.
- D'Aquila, R. T., Bechtel, L. J., Videler, J. A., Eron, J. J., Gorczyca, P., & Kaplan, J. C., (1991). Maximizing sensitivity and specificity of PCR by pre-amplification heating. *Nucleic Acids Research*, 19(13), 3749.
- De Marchi, M., Dalvit, C., Targhetta, C., & Cassandro, M., (2006). Assessing genetic diversity in indigenous Veneto chicken breeds using AFLP markers. *Animal Genetics*, 37(2), 101–105.
- Dieffenbach, C. W., Lowe, T. M., & Dveksler, G. S., (1993). General concepts for PCR primer design. *PCR Methods Appl.*, 3(3), S30–S37.
- Edwards, M. C., & Gibbs, R. A., (1994). Multiplex PCR: Advantages, development, and applications. *Genome Research*, 3(4), S65–S75.
- Espy, M. J., Patel, R., Paya, C. V., & Smith, T. F., (2001). Quantification of Epstein-Barr virus (EBV) viral load in transplant patients by light cycler PCR. In: *101st General Meeting of the American Society for Microbiology*, No. C-148. Orlando, Florida.
- Ferré, F., Marchese, A., Pezzoli, P., Griffin, S., Buxton, E., & Boyer, V., (1994). Quantitative PCR: An overview. *The Polymerase Chain Reaction*, 67–88.
- Freeman, W. M., Walker, S. J., & Vrana, K. E., (1999). Quantitative RT-PCR: Pitfalls and potential. *Biotechniques*, 26(1), 112–125.
- Frey, U. H., Bachmann, H. S., Peters, J., & Siffert, W., (2008). PCR-amplification of GC-rich regions: 'slowdown PCR'. *Nature Protocols*, 3(8), 1312–1317.
- Germer, S., & Higuchi, R., (1999). Single-tube genotyping without oligonucleotide probes. *Genome Research*, 9(1), 72–78.
- Gibson, U. E., Heid, C. A., & Williams, P. M., (1996). A novel method for real time quantitative RT-PCR. *Genome Research*, 6(10), 995–1001.
- Ginzinger, D. G., (2002). Gene quantification using real-time quantitative PCR: An emerging technology hits the mainstream. *Experimental Hematology*, 30(6), 503–512.
- Giulietti, A., Overbergh, L., Valckx, D., Decallonne, B., Bouillon, R., & Mathieu, C., (2001). An overview of real-time quantitative PCR: Applications to quantify cytokine gene expression. *Methods*, 25(4), 386–401.
- Goodman, M. F., (1995). DNA polymerase fidelity: Misinsertions and mismatched extensions. In: *PCR Strategies* (pp. 17–31). Academic Press.

- Green, M. R., & Sambrook, J., (2019). Nested polymerase chain reaction (PCR). *Cold Spring Harbor Protocols*, 2019(2), pdb-prot095182.
- Griffin, H. G., Griffin, A., Nolan, T., & Bustin, S. A., (1994). *PCR Technology: Current Innovations*. CRC Press.
- Hill, D. A., O'Sullivan, M. J., Zhu, X., Vollmer, R. T., Humphrey, P. A., Dehner, L. P., & Pfeifer, J. D., (2002). Practical application of molecular genetic testing as an aid to the surgical pathologic diagnosis of sarcomas: A prospective study. *The American Journal of Surgical Pathology*, 26(8), 965–977.
- Holland, P. M., Abramson, R. D., Watson, R., & Gelfand, D. H., (1991). Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proceedings of the National Academy of Sciences*, 88(16), 7276–7280.
- Innis, M. A., Gelfand, D. H., & Sninsky, J. J., (1999). *PCR Applications: Protocols for Functional Genomics*. Academic Press.
- Innis, M. A., Gelfand, D. H., Sninsky, J. J., & White, T. J., (2012). *PCR Protocols: A Guide to Methods and Applications*. Academic Press.
- International Chicken Polymorphism Map Consortium, (2004). A genetic variation map for chicken with 2.8 million single nucleotide polymorphisms. *Nature*, 432(7018), 717.
- Jarne, P., & Lagoda, P. J., (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution*, 11(10), 424–429.
- Kadri, K., (2019). Polymerase chain reaction (PCR): Principle and applications. In: *Synthetic Biology-New Interdisciplinary Science*. IntechOpen.
- Kemp, D. J., Smith, D. B., Foote, S. J., Samaras, N., & Peterson, M. G., (1989). Colorimetric detection of specific DNA segments amplified by polymerase chain reactions. *Proceedings of the National Academy of Sciences*, 86(7), 2423–2427.
- Komminoth, P., Heitz, P. U., & Long, A. A., (1994). In situ polymerase chain reaction: General methodology and recent advances. *Negotiations of the German Society for Pathology*, 78, 146–152.
- Kwok, P. Y., & Chen, X., (2003). Detection of single nucleotide polymorphisms. *Current Issues in Molecular Biology*, 5(2), 43–60.
- Lawyer, F. C., Stoffel, S., Saiki, R. K., Myambo, K., Drummond, R., & Gelfand, D. H., (1989). Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *Journal of Biological Chemistry*, 264(11), 6427–6437.
- Lay, M. J., & Wittwer, C. T., (1997). Real-time fluorescence genotyping of factor V Leiden during rapid-cycle PCR. *Clinical Chemistry*, 43(12), 2262–2267.
- Lee, L. G., Connell, C. R., & Bloch, W., (1993). Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nucleic Acids Research*, 21(16), 3761–3766.
- Livak, K. J., Flood, S. J., Marmaro, J., Giusti, W., & Deetz, K., (1995). Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *Genome Research*, 4(6), 357–362.
- Lo, Y. D., Wong, I. H., Zhang, J., Tein, M. S., Ng, M. H., & Hjelm, N. M., (1999). Quantitative analysis of aberrant p16 methylation using real-time quantitative methylation-specific polymerase chain reaction. *Cancer Research*, 59(16), 3899–3903.
- Markham, A. F., (1993). The polymerase chain reaction: A tool for molecular medicine. *BMJ: British Medical Journal*, 306(6875), 441.

- McPherson, M. J., Quirke, P., & Taylor, G. R., (1991). *PCR: A Practical Approach*. Oxford–New York–Tokyo: IRL Press at Oxford University Press.
- Mercier, J. F., Slater, G. W., & Mayer, P., (2003). Solid phase DNA amplification: A simple Monte Carlo lattice model. *Biophysical Journal*, 85(4), 2075–2086.
- Mirasena, S., Shimbhu, D., Sanguansermisri, M., & Sanguansermisri, T., (2008). Detection of β -thalassemia mutations using a multiplex amplification refractory mutation system assay. *Hemoglobin*, 32(4), 403–409.
- Mullis, K. B., & Faloona, F. A., (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods in Enzymology*, 155, 335–350.
- Myakishev, M. V., Khripin, Y., Hu, S., & Hamer, D. H., (2001). High-throughput SNP genotyping by allele-specific PCR with universal energy-transfer-labeled primers. *Genome Research*, 11(1), 163–169.
- Negrini, R., Milanese, E., Bozzi, R., Pellicchia, M., & Ajmone-Marsan, P., (2006). Tuscany autochthonous cattle breeds: An original genetic resource investigated by AFLP markers. *Journal of Animal Breeding and Genetics*, 123(1), 10–16.
- Newton, C. R., Graham, A., Heptinstall, L. E., Powell, S. J., Summers, C., Kalsheker, N., & Markham, A. F., (1989). Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Research*, 17(7), 2503–2516.
- Nielsen, R., & Signorovitch, J., (2003). Correcting for ascertainment biases when analyzing SNP data: Applications to the estimation of linkage disequilibrium. *Theoretical Population Biology*, 63(3), 245–255.
- Nijman, I. J., Otsen, M., Verkaar, E. L. C., De Ruijter, C., Hanekamp, E., Ochieng, J. W., & Lenstra, J. A., (2003). Hybridization of banteng (*Bos javanicus*) and zebu (*Bos indicus*) revealed by mitochondrial DNA, satellite DNA, AFLP and microsatellites. *Heredity*, 90(1), 10–16.
- Olek, A., Oswald, J., & Walter, J., (1996). A modified and improved method for bisulphite based cytosine methylation analysis. *Nucleic Acids Research*, 24(24), 5064–5066.
- Osterrieder, N., Hübert, P. H., Brandmüller, C., & Kaaden, O. R., (1994). A touchdown PCR for the differentiation of equine herpesvirus type 1 (EHV-1) field strains from the modified live vaccine strain RacH. *Journal of Virological Methods*, 50(1–3), 129–136.
- Paun, O., & Schönswetter, P., (2012). Amplified fragment length polymorphism: An invaluable fingerprinting technique for genomic, transcriptomic, and epigenetic studies. In: *Plant DNA Fingerprinting and Barcoding* (pp. 75–87). Humana Press.
- PCR Optimization: Reaction Conditions and Components, (2017). *Applied Biosystems*. Available from: https://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_042520.pdf (accessed on 12 October 2022).
- Polymerase Chain Reaction (PCR). National Center for Biotechnology Information [Online]. Available from: <http://www.ncbi.nlm.nih.gov/probe/docs/techpcr/> (accessed on 12 October 2022).
- Primer Design Tips and Tools, (2015). Thermo Fisher Scientific, Inc. Available from: <http://www.thermofisher.com/ca/en/home/products-and-services/product-types/primers-oligosnucleotides/invitrogencustom-dna-oligos/primer-design-tools.html> (accessed on 12 October 2022).
- Rajeevan, M. S., Vernon, S. D., Taysavang, N., & Unger, E. R., (2001). Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. *The Journal of Molecular Diagnostics*, 3(1), 26–31.

- Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlén, M., & Nyrén, P., (1996). Real-time DNA sequencing using detection of pyrophosphate release. *Analytical Biochemistry*, 242(1), 84–89.
- Roux, K. H., (1994). Using mismatched primer-template pairs in touchdown PCR. *BioTechniques*, 16(5), 812–814.
- Roux, K. H., (1995). Optimization and troubleshooting in PCR. *Genome Research*, 4(5), S185–S194.
- SanCristobal, M., Chevalet, C., Haley, C. S., Joosten, R., Rattink, A. P., Harlizius, B., & Cardellino, R., (2006). Genetic diversity within and between European pig breeds using microsatellite markers. *Animal Genetics*, 37(3), 189–198.
- Sapkota, B. R., Ranjit, C., Neupane, K. D., & Macdonald, M., (2008). Development and evaluation of a novel multiple-primer PCR amplification refractory mutation system for the rapid detection of mutations conferring rifampicin resistance in codon 425 of the *rpoB* gene of *Mycobacterium leprae*. *Journal of Medical Microbiology*, 57(2), 179–184.
- Schmittgen, T. D., & Livak, K. J., (2008). Analyzing real-time PCR data by the comparative CT method. *Nature Protocols*, 3(6), 1101–1108.
- Syvänen, A. C., (2001). Accessing genetic variation: Genotyping single nucleotide polymorphisms. *Nature Reviews Genetics*, 2(12), 930–942.
- Teh, L. K., Lee, W. L., Amir, J., Salleh, M. Z., & Ismail, R., (2007). Single step PCR for detection of allelic variation of MDR1 gene (P-glycoprotein) among three ethnic groups in Malaysia. *Journal of Clinical Pharmacy and Therapeutics*, 32(3), 313–319.
- van Huijsduijnen R. H., Ayala, G., & DeLamar, J. F., (1992). A means to reduce the complexity of oligonucleotides encoding degenerate peptides. *Nucleic Acids Research*, 20(4), 919.
- van Pelt-Verkuil, E., Van, B. A., & Hays, J. P., (2008). A brief comparison between *in vivo* DNA replication and *in vitro* PCR amplification. *Principles and Technical Aspects of PCR Amplification*, 9–15.
- Wages, Jr. J. M., (2005). Polymerase chain reaction. *Encyclopedia of Analytical Science*, 243.
- Wilhelm, J., & Pingoud, A., (2003). Real-time polymerase chain reaction. *ChemBiochem*, 4(11), 1120–1128.
- Zhang, X. Y., & Ehrlich, M., (1994). Detection and quantitation of low numbers of chromosomes containing *bcl-2* oncogene translocations using semi-nested PCR. *BioTechniques*, 16(3), 502–507.

CHAPTER 7

Concept and Nature of Genes

ANJUM SABBA¹ and NAJEEBUL TARFEEN²

¹Department of Biochemistry, University of Kashmir, Hazratbal, Srinagar, Jammu and Kashmir, India

²Center for Research and Development (CORD), University of Kashmir, Hazratbal, Srinagar, Jammu and Kashmir, India

ABSTRACT

The concept of gene from mere abstract to a more complex and concrete form has been fascinating through ages now. Its precise definition has not been finalized yet by the biologists because of its dynamic characteristics and as such its deciphering continues till date. The chapter provides a glimpse of the chronological events that were taken forward for proper understanding of the concept of what basically gets transmitted from parent to the offspring in the form of gene and what it is made of, keeping in view its universal mechanism of operation which seems to be same for all living organisms with some minor variations.

7.1 INTRODUCTION

The word “gene” in biology has been derived from the Greek word meaning ‘generation’ (Wikipedia) and is comprehended as trait determinant of an organism (Millard Suman). The journey of gene has started as crude, abstract form to concrete and meaningful entity and is still considered to

be an evolving material and as such its precise definition is also not a static one (Peter Portin). To understand the concept of gene from simple units or elements to more complex and complicated forms of structure, one needs to go back to its historical background to comprehend it in a better possible way. A chronological sequence of events most probably will make it clear as to how much it took to make this word gene from simple units or elements to what we today know as a really complex biological store of information which can be harnessed in different ways. One of the facts of life is that living organisms continue their race by transmitting their characteristics to their offspring in a variety of ways. Some organisms transmit by budding, fragmentation, and spore formation, etc. (Sarin, 2001). But the most common way of transmission of hereditary to the offspring is by sexual reproduction. In all modes of transmission, the parent body, or the parts of the parent body, i.e., gametes are the physical link between the parents and their offspring. This link is the key which contains information from one generation to the subsequent generations and produce same kind yet with a variety of differences within the progeny. Mankind has been known to use selective breeding for improving quality of domestic plants and animals but sometimes the unexpected reoccurrence of undesirable traits could not be explained at that time and it needed careful analysis of transmission of inheritance pattern in progeny (Millard Suman). One such initial insights were provided by Austrian monk Gregor Johan Mendel by his tedious experiments performed in the early 1860's and its deciphering continues till date. This term 'gene' was used in 1909 by Johannsen and still continues to be used till date (Johannsen, 1909). Both genetic as well as biochemical studies corroborated for the mechanism of transmittance of genetic material which makes us to study genes in two broad directions, concept, and the nature of genes, respectively (Demerec, 1961). The events/history of development of concept of gene is too large and for convenience Peter Portin divided it into three phases namely classical, neoclassical, and modern phase.

7.2 CONCEPT OF GENE

7.2.1 CLASSICAL PHASE

The first reported work on what we today know as gene dates back to 1866, when an Austrian monk Gregor Johan Mendel worked out successfully

on some tedious garden pea experiments and hypothesized that there are some genetic factors/determinants/elements which are responsible for the transmittance of traits from parents to offsprings. He used the term factors/elements but not the term gene/allele himself. He gave us an abstract idea of some material present without giving any information regarding its structure or nature. He did his work meticulously and formulated certain laws known now as Mendel's laws of inheritance. One of the conclusions was that if a true-breeding plant with a two contrasting traits is crossed and the progeny selfed, the trait that seemed to be lost, reappeared and the two variants of the trait are assorted independently of each other. His work was not recognized as a big revelation during his life time, though it was quoted 10 times (Gustafson, 1969). It was 16 years after his death that three scientists (de Vries, Correns, and Tschmark) working in different countries (Holland, Germany, and Austria) on different plants and independent of each other came to the same conclusion that were similar to what was discovered by Mendel before 34 years in his publication regarding segregation and independent assortment (de Vries, 1900; Correns, 1900; Tschmark, 1900). That is why it is being called rediscovery of Mendel's laws and therefore, Mendel is regarded as the father of genetics. Mendel suggested the patterns of inheritance he noticed in terms of 'character pairs.' After crossing, these pairs segregate to form 'character pairs' of next generation and independently maintain the behavior irrespective of other character pairs (Hardison, 2020). It was in 1909 that the word gene was coined for the first time by Johannsen and within no time it became the basic unit of genetics and molecular biology. Johannsen regarded it as a mere calculating unit with no emphasis on its nature or physical properties.

The concrete aspect of Mendel's abstract something/factors was provided by two biologists T. H. Boveri and W. S. Sutton working on *Ascaris* and sea urchin, respectively (Boveri, 1902; Sutton, 1903). They found that chromosomes have their role in transmitting traits and soon was called chromosome theory of inheritance. It stipulated the most wanted explanation for Mendel's segregation and independent assortment phenomena for the inheritance of characters/traits by demonstrating that there is continuity of chromosomes after cell division. The theory was developed precisely by Thomas Hunt Morgan school. Shortly after this theory a new phenomenon was observed known as linkage, in which some genes close by were seen to be inherited together thus violating Mendel's

independent assortment law. This linkage paved a way for extension of chromosome theory of inheritance and it was ascertained that genes on different chromosomes show independent assortment in comparison to the genes on same chromosome. But soon after this Morgan extended it further by stating that linkage can be seen in genes lying on the same chromosome as well, provided they are sufficiently far apart from each other and undergo recombination. In short, genes that are closer to each other tend to be co inherited depending upon the distance between them (Morgan, 1915). Morgan and his school showed with the help of *Drosophila* experiments that linkage and its breakdown can be used to map genes on chromosomes and it was A. Stutvar (Morgan's student) who gave us the first abstract map of the *Drosophila* X chromosome in 1913. The X-ray induced changes like translocations and deletions in chromosomes provided another way for the genetic map which was termed as cytological map (Dobzansky, 1929; Muller & Painter, 1929). It was followed by cytological map of giant chromosomes from salivary glands of *drosophila*. It became convenient to identify, visualize, and locate chromosomes by their inheritance pattern (Painter, 1934; Bridges, 1935, 1938). The physical explanation for partial genetic linkage observed by Morgan was provided by Janssens 'chiasmata' or crossing over points, where the actual exchange of genes takes place. Chiasmata or chi shaped structure was observed as a cross figure at the junction or synapses by Janssens (1909) and hypothesized it to be a fusion point between two chromatids of tetrad. Two theories emerged at that time for explaining chiasmata; one called classical hypothesis which could not explain the genetic recombination as it did not consider it to be a point where actual breakage and reunion of chromosome takes place but considered it merely two chromatids crossing over. The other one known as chiasmata theory which assumed that it is simply a result of parental (paternal and maternal) chromatids of tetrad crossing one another. Peter Potin concluded classical era at this stage where a mere abstract concept has turned into something concrete which has been supported with experimental proofs. An unsolved question that remained after the classical concept was 'how' and 'which' portion of chromosome dictate traits in an organism. It was realized that more sophisticated knowledge about physiological, biochemical, and embryological intervention may lead to some solid information regarding these questions. Muller provided a clue for understanding gene nature more extensively by providing an idea that genes might be having some

influential role in the evolution and genesis of life. This very concept by Muller in 1926 gave us a divergent look of gene where it can have long term tremendous effects.

7.2.2 NEOCLASSICAL PHASE

In neoclassical period a gene was initially treated as a dimensionless entity on chromosome which was later given a chemical specification as DNA. This phase began in 1940 with the basic genetics where genes were regarded as a continuous segment having some length (Oliver, Lewis, green, and green). If genes have length, they might be made of something. Since genes reside on chromosomes, now the question was, what part of chromosome actually is responsible for transmitting traits. Chromosomes are almost made up of proteins and DNA, it was a turning point at that time to know which entity (protein/DNA) in the cell is the genetic material. The problem was solved in 1944 by the Oswald's team, who emphasized that it was the DNA which is regarded as the genetic material (Avery et al., 1944). The *in vitro* cell free experiments in which they extracted the substance which caused transformation made them conclude that it was the DNA that changed the morphology of bacterial cells in Griffith's transformation experiments (Griffith, 1928). This answer was not accepted universally, it was after almost eight years later in 1952 when Hershey and Chase proved experimentally that it is only DNA component of bacteriophage responsible for genetic information (Hershey & Chase, 1952). The most fascinating and most wonderful discovery about DNA theory of inheritance was seen in year 1953–1954 with the elucidation of DNA double helical structure by Watson and crick. At this stage, it is pertinent to mention here that Beadle and Tatum had already hypothesized the idea of one gene one enzyme in 1941 (Beadle & Tatum, 1941). By the end of 1950, a solid connection was established between genes and proteins by demonstrating that genes transfer information in the form of mRNA via transcription, which in turn transfer this information to protein synthesis machinery in the cytoplasm (reviewed by potin/Judson). The cis-trans test of B. Lewin (1951) was utilized by Benzer and Yanofsky to demonstrate that cistron is the fundamental unit of gene function. Similarly, the smallest unit of genetic recombination and mutation was known as recon and muton, respectively. In the classical period, gene definition

was regarded as the basic unit of function, recombination, and mutation, which was disapproved by the work of Benzer, who demonstrated that genes had numerous recons and mutons, even a recon or muton can be a single base pair. The collinearity hypothesis which stated that one polypeptide one cistron rather than one gene one enzyme was established as it was observed that some enzymes consist of multimeric units which can either be homogenic or heterogenic and can be the product of different genes (Benzer, 1957; Yanofsky, 1964). The cracking of universal genetic code was taken to be the end of the neoclassical period. According to Peter potin, Genetic code is the triplet nucleotide sequence which determines the sequence of the polypeptide chain. In short, the neoclassical phase gave us information/route map for one gene, one mRNA and one polypeptide concept or more precisely one cistron one polypeptide concept.

7.2.3 MODERN PHASE

The modern phase deals with further deviations of one gene one cistron (mRNA) one polypeptide concept. It was found that a single gene can generate numerous mRNA's and as such, can become a part of several transcriptional units. Two different ways were shown to be responsible for the production of different mRNAs from a single gene. The first one was due to complex promoters (Schibler & Sierra 1987) where RNA polymerase could bind at different sites to initiate transcription, and such the length of mRNA gets altered, which in turn give rise to different polypeptides. The second one was because of alternate splicing (Leff et al., 1986), which happens when exons get merged but not necessarily with the constitutive order. Instead, some exons are alternately picked up for the formation of mature mRNA and rest are spliced out. It depends on the combination of exons and removal of introns that collectively determines the sequence of the polypeptide chain. In short, some exons are cut out from the mature mRNA in some transcripts and may be retained in others. It was also demonstrated that alternate splicing is both tissue and stage-specific. Therefore, the previous concept regarding one gene one mRNA one polypeptide was now taken in new perspective keeping the collinearity of the gene concept intact. This was followed by the identification of two more processes responsible for the production of different functional products like protein from one gene, i.e., RNA editing and gene

sharing (Brennickle et al., 1999; Piatigorsky, 2007). RNA editing involves the post transcriptional base modification of the transcript, which in turn alters the amino acid sequence and thereby the function of the final product protein. Another way of getting different function from the same gene is what is known as ‘protein moonlighting.’ In this process a gene produces same transcript but after the translation the polypeptide may get different configuration/folding pattern which alters the function of the protein.

At the end of modern phase gene is still regarded as a DNA sequence on chromosome which dictates information for the production of polypeptides and proteins but one gene sequence may be used to harvest different polypeptides which in turn exhibit different functions depending upon the combination of configuration it carries. In short, the natural manipulation of genes to produce different functional products can be observed at different levels like transcriptional, post-transcriptional, translational, and post translational levels. The functional and structural aspect of gene is still under continuous construction as new and more sophisticated knowledge related to DNA adds to its complexity. Some of the most prominent known natural features associated with DNA and which adds to complexities and somehow does not hold true for classical/neoclassical assumption are discussed in subsections.

7.2.3.1 OPEN READING FRAME (ORF)

The three-letter genetic code of a DNA sequence is called codon and each triplet signifies which amino acid is to be incorporated depending upon the triplet. If one of the three bases of a codon gets added or deleted while the rest remains the same, a phenomenon known as frameshift mutation occurs, which makes the entire frame of DNA sequence to shift a notch as compared to its original sequence. Which in turn changes the basic structure of a polypeptide. In contrast, if all the three bases of a single codon are deleted or added in a DNA sequence, only one amino acid change is seen in the polypeptide chain and retains almost all basic structure. Some of the DNA sequences found to have transcription start site with no stop codon popularly known as open reading frames (ORF) which in certain cases are considered to be a gene or a part of it (Hozomi & Tonegawa, 1976) does not fit any of the previous rational definition of the gene.

7.2.3.2 POLY A SITES

During the formation of mature mRNA from heteronuclear RNA, a unique feature has been observed at the 3-prime end of maturing mRNA, i.e., incorporation of about 200 adenine nucleotides. This addition of 200 adenine molecules are not dictated by DNA, but certain sites have been found where poly adenylation can occur within mRNA sequence. This indicates that one mRNA molecule can be polyadenylated at many sites, thereby changing the length of mRNA, which in turn can affect the polypeptide length (Leff et al., 1986). Thus, one gene can give rise to more than one polypeptide which violates classical/neoclassical concept of gene.

7.2.3.3 NESTED GENES

One of the classical concepts of gene was thought to be the linear sequential arrangement of genes on chromosomes, and nested genes are found in contradiction to it. These nested genes, in simple terms, are genes within gene. These genes are found to be within the intron sequence of other gene and was for the first time illustrated by Henikoff et al. (1986) in drosophila and in man by Levinson et al. (1990).

7.2.3.4 PSEUDO GENES

These are the DNA sequences that resemble functional gene but possesses a change in such a way so that it will not perform the normal function Reiger et al. (1991). It was first observed in *Xenopus laevis* by Jacq et al. (1997). Plenty of multigene families in high up eukaryotes are found to possess pseudogenes. They are thought to be ancestral reading frames which has been closed during evolution due to the incorporation of many stop codons.

7.2.3.5 MOVABLE GENES

Movable genes also known as insertion sequences/transposons, and as such are the stretches of DNA which can shift from one site to another in genome. It was McClintock (1947, 1948) who first of all noticed this movement of DNA sequences which she named control elements. These sequences have the ability to move from one chromosome, individual or species to another chromosome, organism, and species, respectively. Since

the location of the gene is not fixed, it does not go with the classical/neoclassical hypothesis of fixed location of genes on the chromosomes.

7.2.3.6 REPEATED GENES

Repetitive DNA was first demonstrated by Waring/Britten and Britten/Kohne (1966–1968). Reassociation kinetic studies done on DNA of various organisms like amphibian oocytes Miller & Beatty (1969), sea urchin Weinberg et al. (1972), bacteria, drosophila, avian, mammalian cells and higher eukaryotes (reviewed in Benjamin 1980a) have revealed the presence of repetitive DNA. According to Dillon gene organization with respect to DNA repetition can be classified into five categories:

- i. Single copy genes, i.e., simplest genes with no DNA repetition sequence;
- ii. Genes whose few to many copies are found scattered in the genome of the organism;
- iii. Genes whose sequence is repeated one or few times on the same strand;
- iv. Genes Which are in multiple copies and the gene family is scattered in the genome; and
- v. Genes which form clusters and are repeated in tandem manner.

The genes which are repeated on the same strand or in tandem repeats are basically transmitted as blocks, and it is this DNA repetition that goes against the classical concept, as the gene is not transcribed as a single transcription unit. In other words, the unit for transmission and transcription are not identical.

7.2.3.7 POLYPROTEIN GENES

When proteolytic cleavage of a single translated product pertaining to one gene yields many polypeptides, they are known as polyprotein genes. Neuropeptides Douglas et al. (1984) and human salivary proline-rich proteins Lyons et al. (1988) are the well-known examples of poly protein genes. One gene—one polypeptide concept of neoclassical concept is violated as far as polyprotein gene is concerned.

7.2.3.8 *ENHANCERS/SILENCERS*

Enhancers are DNA elements located usually upstream or downstream in the 5' and 3' location of a gene but can be found within the gene in introns and exons (Nikovits et al., 1990) also. The first enhancer element was found in the SV40 virus and the first cell specific enhancer was found in heavy chain of immunoglobulin gene by Banerji et al. (1981, 1983) and Gillies et al (1983) respectively. Enhancers can regulate gene positively or negatively (silencers). They can act in cell specific, tissue-specific and host specific manner. Their action may be constitutive or can be stimulated by response to some inducible signal also. In view of neoclassical concept of gene, enhancers provide another level of complexity and therefore must be considered as a part of gene even though they sometimes act within and outside the gene transcript, being cis acting elements.

7.2.3.9 *ASSEMBLED GENES/IMMUNOGLOBULIN GENES*

The enormous diversity found in the production of antibodies from the genome was long considered to be a question of concern for genetics. The answer was provided by the phenomenon – somatic recombination of immunoglobulin genes, in which there occurs an internal rearrangement/combination and permutation of different regions of variable and constant portion of the concerned immunoglobulin gene. It means that the gene information in germline and mature immune cells differ completely, which does not abide by the classical or neoclassical concept of gene (Hozumi & Tonegawa, 1976).

7.3 **THE CRUX AND CRACKS IN THE CONCEPT OF GENE**

Recent developments in the advancement of technology and techniques have made our understanding regarding gene concept to be taken in a modified version. The classical and neoclassical concept of gene has to be viewed in different and more concrete perspective, taking into consideration the complexities which have been observed since its inception. Some of the observed complexity is mentioned below:

- The gene in eukaryotes does not usually possess a perfect boundary of start and stop for transcripts. In reality, the whole chromosome seems to be a continuous stretch for different transcriptions, making it more difficult to interpret its one-to-one relationship between primary transcript and its final product.
- Human genome can transcribe both protein-coding and noncoding regions from both the strands of DNA, which may be overlapped and interlaced.
- There are transcripts known to have protein exons combined from different locations of the genome hundreds of thousands nucleotide away. Similarly, exons of different genes have been observed to be a part of a single transcript, most probably known as gene fusion (Parra et al., 2006), in comparison to some microbial eukaryotes where encrypted genes are found as separate fragments of its genome. In some multicellular eukaryotes different parts of non-homologous chromosomes have also been seen in protein-coding transcripts.
- The epigenetic inheritance provided another insight into the complexity of gene. It does not hold true for the classical/neoclassical concept of gene as the functional quality of some genes are inherited to future generations as well.
- DNA sometimes can restore its ancestral form by repairing its mutant site with the help of template directed RNA cache.
- Almost all genes form transcripts which most often yield protein as a final product yet a lot of RNA encoding genes have been seen engaged in transcription of a variety of RNA transcripts only known as non-coding RNA, which may be small like siRNA, miRNA, shRNA, and long non-coding RNA (lncRNA). Recent studies have shown their critical role in the regulation of gene and therefore serve some biological function which makes them eligible to be called genes.

7.4 NATURE OF GENE

The reaffirmation of work done by the father of genetics, G.J. Mendel was done by H. Vries, C. Correns, and E. Tschermak. ‘Characters’ were considered as the expressions of hereditary particles in a cell. It was W

Bateson who named these hereditary particles as ‘allelomorphs’ or ‘alleles’ (Hardison, 2020). The mechanism of transmittance of genetic material is assumed to be the same in all living organisms, i.e., from the smallest known microorganism, bacteriophages to the most complex mammals and the evidence for the same have been derived from both genetic research and biochemical studies. Despite, fundamental similarity in all living forms, some significant variations have been seen in the form of RNA/DNA, single stranded DNA/double stranded DNA being genetic material in some plant/animal viruses/coliphages/S-13 phage and other organisms, respectively. The nature of genes can be understood keeping in view the universal mechanism of transmittance, considering both genetic (physical properties) and biochemical aspect as well (Demerec, 1961).

7.4.1 PHYSICAL NATURE OF GENE

Genes perform their function by directing the protein synthesis. In Eukaryotes, the nucleus is the site where almost all genes are located. Besides, a small proportion of genes in plants and animals are also harbored by chloroplasts and mitochondria respectively that have small subsets of genes within chloroplasts and mitochondria, respectively, and are distinct from nuclear genes. In bacteria, genes are present on the chromosomes that are floating within the cytoplasm. The bacteria may also contain extrachromosomal double circular DNA elements known as plasmids where the genes are much less in number. In an organism, genome content comprises of a number of genes that varies between species. For instance, the genome of an *Escherichia coli* (*E. coli*) houses approximately 5,416 genes, whereas, the genome of humans contains almost 20,000 to 25,000 genes. In case of *Arabidopsis thaliana*, the number of genes recovered was roughly 25,500. Among extant organisms, *Mycoplasma genitalium* has approximately 517 number of genes (Leslie, 2008).

7.4.1.1 PHYSICAL LENGTH OF A GENE

The genes range in size from 1,148 to 37.7 kb for humans and for mouse the size ranges from 994 to 41.8 kb. The number of exons varies from 1 to 41 per gene on an average of 8.4 (Pearson et al., 1988). Various studies suggested that the conserved genes are having more intronic content and

are longer in length (Vishnoi et al., 2010). While as gene with shorter length have little intronic burden, high expression, and smaller proteins (Urrutia & Hurst, 2003). It is hypothesized that, because of higher levels of expression in smaller genes, there occurs a selective pressure to increase the efficiency of protein synthesis (Urrutia & Hurst, 2003). In humans, longer length genes, are present in brain, muscles, and are important in early developmental stages and there is a probability of high rates of mutations (Helmrich et al., 2011; Zylka et al., 2015; Takeuchi et al., 2018; Hosokawa et al., 2019). Smaller gene groups are mostly associated with skin including cornified envelope and skin development (Pipkin & Monticell, 2008); stomach, testis, pancreas, and immune system (Lopes et al., 2021). Studies showed that genes, having long transcripts have an excessive number of co-expressed genes and protein-protein interactions (Lopes et al., 2021).

It has been estimated that in humans there occurs the protein-encoding genes that are approximately 20,000–25,000 in number, In *Arabidopsis thaliana*, number is 25,000 and is considered one of the plants with smallest genome content in plant kingdom. In *Caenorhabditis elegans* and *Drosophila melanogaster*, the number of proteins encoding genes are 18,000 and 14,000, respectively.

7.4.1.2 GENE LOCI

In the initial decade of the 20th century, it was not clear that segregation and independent assortment of alleles could provide an explanation of patterns of inheritance. Karl Pearson and Walter F. R. Weldon, who were the contemporaries of Bateson, explained that a gene is referred to as a unit of genetic information, that is heritable and is located at a particular fixed spot known as locus on a chromosome (Hardison, 2020). Genetists have constructed the maps that showed the relative positions of genes in the chromosomes of various organisms, by observing the recombination frequencies between genetic markers positioned on each chromosome. The same principle is now considered to study the structure of gene and to work out maps of various gene loci. Gene loci consist of linearly arranged sites or subunits, and site is a subunit of locus at which there is a possibility of mutation and within which recombination presumably does not occur. Evidence showed that gene locus contains a large number of sites. In T4

phage, in the rII locus at least 400 sites had been identified by Benzer, with no indication that the saturation point has been approached (Benzer, 1955). In Salmonella, Leu locus contains 52 different sites without any indication of mutation at the same site. Levine showed that there occurs 2000 number of sites in that locus (Glanville & Demerec, 1960).

The list of loci arranged in an ordered manner for a particular genome is termed as a gene map. Gene mapping is a method of determining the particular locus or loci responsible for producing a specific biological trait. Early gene maps used the linkage analysis. In a chromosome, more the two genes are close to each other, more is the chances of their inheritance together. Two distinct maps, one physical map and the other genetic map have been used to study linkage phenomenon. Physical maps measure the number of base pairs or physical distance between base pairs by using markers while as, genetic map tells us about the nature of different regions of chromosome (Alizadeh et al., 1995; Griffiths, 2000).

7.4.1.3 CONCEPT OF ALLELISM

Allele is referred to as different variants of a gene at a particular locus. Various genetists have defined the criteria of differentiating between allelic mutants. These comprise of degree of mutability, either spontaneous or induced, pattern of mutability, capability to grow on minimal medium or potential to grow at variable temperatures, complementation relations and response to the suppressor mutations. Mutation at the same site may also leads to different alleles that are phenotypically different. The sites of alleles that are responsible for related phenotypes are distributed along a gene locus (Demerec, 1961). Chromosomes have the same alleles at a specific locus in diploid and polyploidy cells, and these are referred to as homozygous, while those with different alleles are referred to as heterozygous.

7.4.1.4 BOUNDARIES OF THE LOCUS

Certain genes are grouped together that have related functions and are organized in the same order as a chain of biosynthetic reactions they control. This type of arrangement allows us to find out the degree of certainty, when two loci are in close proximity to one another and to study

the areas close to their boundary line. For instance, two loci *hisB* and *hisC* that were studied by Hartman revealed that 31 sites containing 34 alleles of *hisB* locus regulates the enzyme imidazole glycerol phosphate ester dehydrase and *hisC* locus containing 33 sites with 35 alleles regulates the imidazole acetol phosphate ester transaminase suggested the demarcation line between the B and C loci (Demerec & Hartman, 1959).

7.4.2 CHEMICAL NATURE OF GENE

Chemically genes are mostly composed of deoxyribonucleic acid (DNA), but in some like viruses, it is made of ribonucleic acid (RNA). It was investigated by various biologists that confirmed that chemically genes are made up of DNA. It was in 1923, Frederick Griffith, in his transformation experiment revealed that DNA acts as genetic material by observing III-S strain, in which DNA contains the genes that forms smooth polysaccharide capsule which was later on fully confirmed by O. Avery, a Canadian born American bacteriologist C. Macleod, a Canadian American genetist and M. McCarty, an American genetist (Avery et al., 1944). In 1953 James D Watson and Francis H Crick were the first to show that the nucleotides forming the two chains twist around one another to form a twisted ladder (Watson & Crick, 1953). The two chains of DNA are having opposite polarity (antiparallel) and runs in an opposite direction. The two DNA strands are termed as polynucleotides made up of monomeric units (Alberts et al., 2002). The two side chains of a ladder are composed of phosphate and deoxyribose sugar, and nitrogenous bases form the steps of a ladder (Ghosh & Bansal, 2003). The nucleotides are connected to one another by a phosphodiester bond that exists between ribose of one nucleotide and attached phosphate of successive nucleotides which results in sugar-phosphate outer backbone. Phosphodiester occurs between the third and fifth carbon atoms of adjacent deoxyribose sugar. These are termed as 3' and 5' end carbons (Berg et al., 2002). The nitrogenous bases are of two types purines and pyrimidines, Purines can be adenine (A), guanine (G) and pyrimidines can be cytosine (C) and thymine (T) (Sanger, 1984). The A of one side of the chain bonds to the T of the other, while as, C on the one string links to the G on the other side forming the rungs of the ladder. If bonds are broken between the bases, the two chains of ladder are untwisted, and free nucleotides attach themselves to the free bases of

the now-separated chains. Stability of DNA relies on the presence of high content of GC in the Polynucleotide chain. Studies by Erwin Chargaff discovered that content of thymine is equal to that of cytosine. These base ratios led to the base-pairing rules, which were the foundations that led to the discovery of the double helix. The nucleotides that are free arrange themselves in each chain in accordance with the base-pairing rule where A bonds to T and C bonds to G. Base pairing involves the formation of hydrogen bonds in which A forms double bond with T and G forms a triple bond with C (Troppe, 2012). This process leads to the formation of two similar DNA molecules from parent/original one and is recognized as a way of transferring the blue print from parent to their offspring or from one generation to the other. The structure of gene comprises of many components that not only includes protein coding sequence but also DNA regions that are not transcribed, and also contains the untranslated part of RNA. Genes contain regulatory sequences that have a role in increasing or decreasing the expression of specific genes. Genes also have a promoter to which transcription factors bind and also aids in the recruitment of RNA polymerase to initiate transcription (Alberts et al., 2002). Promoters are the short DNA stretches on chromosome found upstream of the start site of transcription and can be 100–1,000 bp long. Many eukaryotic genes contain ‘TATA box’ as their promoter sequence that are located at ‘25 to 35 bp’ upstream of the transcription initiation site and are conserved in nature. An initiator element occurs in between –3 and +5 of the regions. Other promoter elements are found in different promoter regions of eukaryotes. The promoter region in the case of prokaryotes comprises of two conserved sequences that are positioned at ‘-10 and -35 ahead’ from the transcription initiation site. The ‘-10 element’ sequence is known as ‘Pribnow box’ and consists of ‘TATAAT’ sequence while as, other sequence is ‘TTGACA’ and is positioned at ‘-35 element’ site. The sequence at –10 is necessary to initiate transcription in prokaryotes and sequence at -35 allows very high transcription rate (Pierce & Benjamin, 2002; Vassilyev et al., 2002). A gene can have multiple promoters, that results in variants of mRNAs obtained from a single gene that corresponds to how long they extend in 5’-end (Mortazavi et al., 2008). Strong promoter sequences are found in highly transcribed genes, and they generate a strong interaction with transcription factors, causing mRNA to be formed more frequently. In contrast, promoters in other attenuated genes create weak associations with transcription proteins or factors, resulting in a low rate of mRNA

synthesis. When the promoter regions of eukaryotes and prokaryotes were compared, eukaryotic promoters were found more complex and harder to be recognized (Alberts et al., 2002). Besides that, there are regulatory areas in genes that are placed many kilobases downstream/upstream of the (ORF) that have a role in altering the gene expression. These regulatory sites act by associating with factors necessary for initiation of transcription that makes the DNA to bend in the form of a loop and causes the transcription factors bound to regulatory sequences to act as a complex at the site where RNA polymerase binds and initiates transcription in a more efficient way. Enhancer is basically a short stretch of DNA that strongly stimulates transcription of transcription unit. An enhancer can stimulate transcription over long distances as many as 1,000 base pairs and acts on downstream or upstream of the transcription initiation site (Smemo et al., 2014). Enhancer in association with activator proteins acts on the promoter and recruits general factors of transcription and RNA polymerase which then initiates gene transcription (Dong et al., 2010; Birnbaum et al., 2012). The transcribed pre-mRNA at both ends contains untranslated regions (5'UTR and 3'UTR). Following 5'UTR there occurs initial exon which contains start codon, followed by series of introns and exons alternatively which then ends with terminating exon containing stop codon, it is then followed by 3'UTR at its end (Mignone et al., 2002). Additionally, ORFs in most eukaryotes contains untranslated introns that are first removed prior to exon translation. The ends of the introns contain sequences that have a role in splicing so to produce mature mRNA. Eukaryotic mRNA precursors are processed to eliminate introns and adding of caps at 5 prime end and poly(adenine) sequences at 3 prime end is carried out to produce functional mRNA. Usually introns start with 'GT' sequence (GU in RNA) and terminates with 'AG' sequence. Intron sequences may be large comparative to coding sequences, over 90% of sequence in some genes that lies between 5' and 3' ends of mRNA are intronic in nature. Eukaryotic genes may be clustered but are controlled independently (Guhaniyog et al., 2001). In contrast to prokaryotes where genes are sorted into operons, having numerous sequences that are coding in nature, and these coding sequences are transcribed as a unit (Salgado et al., 2000; Blumenthal, 2004). In prokaryotes where a single gene often controls a whole cluster (for instance, LacI regulates the synthesis of three enzymes β -galactosidase, permease, and acetylase which helps in the degradation of lactose). Thus, the genes in an operon forms a continuous

mRNA that is known as polycistronic mRNA (Clean & Phillip, 1997). The term cistron is an alternate term for 'gene.' In an operon, transcription is often controlled by a repressor, and it depends on the presence of specific metabolites, which in turn depends on whether the repressor is present in an inactive or an active state. The operon genes that are transcribed into proteins have similar functions and are engaged in the same regulatory network (Bruce et al., 2002).

KEYWORDS

- **deoxyribonucleic acid**
- **gene**
- **long non-coding RNA**
- **molecular biology**
- **open reading frames**
- **ribonucleic acid**

REFERENCES

- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P., (2002). How genetic switches work. *Molecular Biology of the Cell* (4th edn.).
- Alizadeh, F., Karp, R., Newberg, L., & Weissner, D., (1995). Physical mapping of chromosomes: A combinatorial problem in molecular biology. *Algorithmica*, 13, 52–76.
- Avery, O. T., MacLeod, C. M., & McCarty, M., (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a deoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *J. Exp. Med.*, 79, 137–159.
- Banerji, J., et al., (1981). Expression of a globin gene is enhanced by remote SV40 DNA sequences. *Cell*, 27, 299–308.
- Banerji, J., et al., (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell*, 33, 729–740.
- Beadle, G. W., & Tatum, E. L., (1941). Genetic control of biochemical reactions in neurospora. *Proceedings of the National Academy of Science USA*, 27, 499–506.
- Benzer, S., (1955). Fine structure of a genetic region in bacteriophage. *Proc. Nat. Acad. Sc.*, 41, 344–354.
- Benzer, S., (1957). The elementary units of heredity. In: McElroy, W. D., & Glass, B., (eds.), *The Chemical Basis of Heredity* (pp. 70–93). Baltimore: Johns Hopkins Press.

- Berg, J., Tymoczko, J., & Stryer, L., (2002). *Biochemistry*. W.H. Freeman and Company. ISBN 0-7167-4955-6.
- Bibliographic Details for "Gene." *Wikipedia, the Free Encyclopedia*. Page Version ID: 1022654359, Permanent link: <https://en.wikipedia.org/w/index.php?title=Gene&oldid=1022654359>.
- Birnbaum, R. Y., Clowney, E. J., Agamy, O., et al., (2012). Coding exons function as tissue-specific enhancers of nearby genes. *Genome Res.*, 22(6), 1059–1068.
- Blumenthal, T., Evans, D., Link, C. D., Guffanti, A., Lawson, D., Thierry-Mieg, J., Thierry-Mieg, D., et al., (2002). A global analysis of *Caenorhabditis elegans* operons. *Nature*, 417, 851–885.
- Boveri, T., (1902). Über mehrpolige mitosen als Mittel zur analyse des zellkerns. *Verh. Phys-Med. Ges. Wurzb.* 35, 60–90.
- Brennickle, A., Marchfelder, A., & Binder, S., (1999). RNA editing. *FEMS Microbiol. Rev.*, 23, 297–316.
- Bridges, C. B., (1935). Salivary chromosome maps with a key to the banding of the chromosomes of *Drosophila melanogaster*. *J. Hered.*, 26, 60–64.
- Bridges, C. B., (1938). A revised map of the salivary gland X-chromosome of *Drosophila melanogaster*. *Journal of Heredity*, 29, 11–13.
- Britten, R. J., & Kohne, D. E., (1967). *Nucleotide Sequence Repetition in DNA* (Vol. 65, pp. 78–106). Carnegie Institute Washington Yearbook.
- Britten, R. J., & Kohne, D. E., (1968a). Repeated sequences in DNA. *Science*, 161, 529–540.
- Britten, R. J., & Kohne, D. E., (1968b). Repeated nucleotide sequences. *Carnegie Institute Washington Yearbook*, 66, 73–88.
- Clean, M. P., (1997). *Prokaryotic Gene Expression*. ndsu.edu.
- Correns, C., (1900). G. Mendel's rule on the behavior of the offspring of hybrid races. *Reports of the German Botanical Society*, 18, 158–168.
- De Vries, H., (1900). The law of division of the bastards. *Reports of the German Botanical Society*, 18, 83–90.
- Demerec, M., & Hartman, P. F., (1959). Complex loci in microorganisms. *Ann. Rev. Microbiol.*, 13, 377–406.
- Demerec, M., (1960). *Presented as Part of a Symposium on Genetics Held at Western Reserve University*. The symposium was supported by National Science Foundation Grant G 9370.
- Demerec, M., (1961). The Nature of the gene. *Am J Hum Genet.*, 13(1 Pt 2), 122–127.
- Dobzhansky, T., (1929). Genetical and cytological proof of translocations involving the third and fourth chromosome in *Drosophila melanogaster*. *Biologisches Zentralblatt*, 49, 408–419.
- Dong, X., Navratilova, P., Fredman, D., et al., (2010). Exonic remnants of whole-genome duplication reveal cis-regulatory function of coding exons. *Nucleic Acids Res.*, 38(4), 1071–1085.
- Douglass, J., et al., (1984). Polypeptide gene expression: Generation of diversity of neuroendocrine peptides. *Annual Review of Biochemistry*, 53, 665–715.
- Ghosh, A., & Bansal, M., (2003). "A glossary of DNA structures from A to Z". *Acta Crystallographica Section D.*, 59(Pt 4), 620–626.

- Gillies, S. D., (1983). A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell*, 33, 717–772.
- Glanville, E. V., & Demerec, M., (1960). Threonine, isoleucine, and isoleucine-valine mutants of *Salmonella typhimurium*. *Genetics*, 45, 1360–1374.
- Green, M. M., & Green, K. C., (1949). Crossing over between alleles of the lozenge locus in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, 35, 586–591.
- Griffith, F., (1928). Significance of pneumococcal types. *Journal of Hygiene Cambridge*, 27, 113–159.
- Griffiths, A. J. F., (2000). *An Introduction to Genetic Analysis* (7th edn.). W. H. Freeman.
- Guhaniyogi, J., & Brewer, G., (2001). Regulation of mRNA stability in mammalian cells. *Gene*, 265(1–2), 11–23.
- Gustafsson, A. Ê., (1969). The life of Gregor Johan Mendel – tragic or not? *Hereditas*, 62, 239–258.
- Hardison, R., (2020). *Fundamental Properties of Gene*. Chapter 1; LiberText biology.
- Helmrich, A., Ballarino, M., & Tora, L., (2011). Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol. Cell*, 44, 966–977.
- Henikoff, S., et al., (1986). Gene within a gene: Nested *Drosophila* genes encode unrelated proteins on opposite DNA strands. *Cell*, 44, 33–42.
- Hershey, A. D., & Chase, M., (1952). Independent functions of viral protein and nucleic acid in growth of bacteriophage. *Journal of General Physiology*, 36, 39–56.
- Hosokawa, M., Takeuchi, A., Tanihata, J., Iida, K., Takeda, S., & Hagiwara, M., (2019). Loss of RNA-binding protein Sfpq causes long-gene transcriptopathy in skeletal muscle and severe muscle mass reduction with metabolic myopathy. *iScience* 13, 229–242.
- Hozumi, N., & Tonegawa, S., (1976). Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proceedings of the National Academy of Science USA*, 73, 3628–3632.
- Jacq, C., et al., (1977). A pseudogene structure in 5S DNA of *Xenopus laevis*. *Cell*, 12, 109–120.
- Janssens, F. A., (1909). La théorie de la chiasmatisation, nouvelle interpretation des cinèses de maturation. *Cellule*, 25, 387–406.
- Johannsen, W., (1909). Elements of the exact theory of heredity. *Jena: Gustav Fischer*, 143–144.
- Judson, H. F., (1996). *The Eighth Day of Creation: Markers of the Revolution in Biology*. Expanded Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Leff, S. E., Rosenfeld, M. G., & Evans, R. M., (1986). Complex transcriptional units: Diversity in gene expression by alternative RNA processing. *Annu. Rev. Biochem.*, 55, 1091–1117.
- Leslie, A. P., (2008). Eukaryotic genome complexity. *Nature Education*, 1(1), 96.
- Levinson, B., et al., (1990). A transcribed gene in an intron of the human factor VIII gene. *Genomics*, 7, 1–11.
- Lewin, B., (1980a). *Gene Expression, Eukaryotic Chromosomes* (Vol. 2., 2nd edn.). New York: Wiley.
- Lewis, E. B., (1941). Another case of unequal crossing over in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, 27, 31–34.

- Lewis, E. B., (1945). The relation of repeats to position effect in *Drosophila melanogaster*. *Genetics*, 30, 137–166.
- Lewis, E. B., (1951). Pseudoallelism and gene evolution. *Cold Spring Harb. Symp. Quant. Biol.*, 16, 159–174.
- Lopes, I., Altab, G., Raina, P., & Magalhães, J. P., (2021). Gene size matters: An analysis of gene length in the human genome. *Front. Genet.*, 12, 559998.
- Lyons, K. M., et al., (1988). Many protein products from a few loci: Assignment of human salivary proline-rich proteins to specific loci. *Genetics*, 120, 255–265.
- McClintock, B., (1947). Cytogenetic studies of maize and neurospora. *Carnegie Institute Washington Yearbook*, 46, 146–152.
- McClintock, B., (1948). Mutable loci in maize. *Carnegie Institute Washington Yearbook*, 47, 155–169.
- Mignone, F., Gissi, C., Liuni, S., & Pesole, G., (2002). Untranslated regions of mRNAs. *Genome Biol.*, 3, REVIEWS0004.
- Miller, O. L. Jr., & Beatty, B. R., (1969b). Extrachromosomal nucleolar genes in amphibian oocytes. *Genetics*, 61, 133–143.
- Morgan, T. H., et al., (1915). *The Mechanism of Mendelian Heredity*. New York: Henry Holt.
- Mortazavi, A., et al., (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods*, 5, 621–628.
- Muller, H. J., & Painter, T. S., (1929). The cytological expression of changes in gene alignment produced by X-rays in *Drosophila*. *American Naturalist*, 63, 193–200.
- Muller, H. J., (1926). Gene as the basis of life. *Proc. Internat. Cong. Plant Sci.*, 1, 897–921.
- Nikovits, W. Jr., Mar, J. H., & Ordahl, C. P., (1990). Muscle-specific activity of the skeletal troponin I promoter requires interaction between upstream regulatory sequences and elements contained within the first transcribed exon. *Mol. Cell. Biol.*, 10, 3468–3482.
- Oliver, P., (1940). A reversion to wild type associated with crossing over in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, 26, 452–454.
- Painter, T. S., (1934). A new method for the study of chromosome aberrations and the plotting of chromosome maps in *Drosophila melanogaster*. *Genetics*, 19, 175–188.
- Parra, G., Reymond, A., Dabbousch, N., Dermitzakis, E. T., Castelo, R., et al., (2006). Tandem chimerism as a means to increase protein complexity in the human genome. *Genome Res.*, 16, 37–44.
- Pearson, W. R., & Lipman, D. J., (1988). *Proc. Nat. Acad. Sci., U.S.A.*, 85, 2444–2448.
- Piatigorsky, J., (2007). *Gene Sharing and Evolution: The Diversity of Protein Functions*. Harvard University Press, Cambridge, MA.
- Pierce, B. A., (2002). *Genetics: A Conceptual Approach* (1st edn.). New York: W.H. Freeman and Co.
- Pipkin, M. E., & Monticelli, S., (2008). Genomics and the immune system. *Immunology*, 124, 23–32.
- Portin, P., & Adam, W., (2017). The evolving definition of term “gene”. *Genetics Perspectives*, 205, 1353–1364.
- Portin, P., (1993). The concept of the gene: Short history and present status. *Quarterly Review of Biology*, 68(2), 173–223.
- Portin, P., (2002). Historical development of the concept of the gene. *Journal of Medicine and Philosophy*, 27(3), 257–286.

- Rieger, R., et al., (1991). *Glossary of Genetics: Classical and Molecular* (5th edn.). Berlin and Heidelberg: Springer-Verlag.
- Salgado, H., Moreno-Hagelsieb, G., Smith, T., & Collado-Vides, J., (2000). Operons in *Escherichia coli*: Genomic analyses and predictions. *Proceedings of the National Academy of Sciences*, 97(12), 6652–6657.
- Sanger, W., (1984). *Principles of Nucleic Acid Structure*. New York: Springer-Verlag. ISBN 0-387-90762-9.
- Sarin, C., (2001). *Genetics*. Tata McGraw- Hill. ISBN 0-07-460333-7.
- Schibler, U., & Sierra, F., (1987). Alternative promoters in developmental gene expression. *Annu. Rev. Genet.*, 21, 237–257.
- Smemo, S., Tena, J. J., Kim, K. H., Gamazon, E. R., Sakabe, N. J., Gómez-Marín, C., et al., (2014). Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature*, 507, 371–375.
- Sturtevant, A. H., (1913). The linear arrangement of the six sex-linked factors in *Drosophila*, as shown by their mode of association. *J. Exp. Zool.*, 14, 43–59.
- Suman, M., (2001). *Genes: Definition and Structure*. USA encyclopedia of life sciences; Nature Publishing Group; www.els.net 1 (accessed on 12 October 2022).
- Sutton, W. S., (1903). The chromosomes in heredity. *Biol. Bull.*, 4, 231–251.
- Takeuchi, A., Iida, K., Tsubota, T., Hosokawa, M., Denawa, M., Brown, J. B., et al., (2018). Loss of Sfpq causes long-gene transcriptopathy in the brain. *Cell Rep.*, 23, 1326–1341.
- Tropp, B. E., (2012). *Molecular Biology* (4th edn.). Sudbury, Mass.: Jones and Barlett Learning. ISBN 978-0-7637-8663-2.
- Tschermak, E., (1900). Ueber Ku Ènstliche kreuzung bei *Pisum sativum*. *Berichte der Deutschen Botanischen Gesellschaft*, 18, 232–239.
- Urrutia, A. O., & Hurst, L. D., (2003). The signature of selection mediated by expression on human genes. *Genome Res.*, 13, 2260–2264.
- Vassilyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassilyeva, M. N., Borukhov, S., & Yokoyama, S., (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution *Nature*, 417, 712–719.
- Waring, M., & Britten, R. J., (1966). Nucleotide sequence repetition: A rapidly reassociating fraction of mouse DNA. *Science*, 154, 791–794.
- Watson, J. D., & Crick, F. H. C., (1953). A structure for deoxyribonucleic acid. *Nature*, 171, 737–738.
- Weinberg, E. S., et al., (1972). Genes coding for polysomal 9S RNA of sea urchins: Conservation and divergence. *Nature*, 240, 225–228.
- Yanofsky, C., et al., (1964). On the colinearity of gene structure and protein structure. *Proceedings of the National Academy of Science USA*, 51, 266–272.

CHAPTER 8

Blotting Techniques

ASIMA TAYYEB, ZHUHA BASIT, and HANFA ASHFAQ

*School of Biological Sciences, University of the Punjab, Lahore,
Pakistan*

ABSTRACT

Identification of mutated and abnormal genes is an important goal in genetic counseling and clinical research. Identification and measurement of specific proteins in biological sample is quite laborious and complex. Blotting techniques are gaining importance for identification of unique nucleic acid and protein sequences. Their role is increasing in clinical research and molecular biology.

8.1 GENERAL PRINCIPLE

The blotting techniques are quite simple and there are four steps in general. Separation of nucleic acids or proteins through electrophoresis, the separated nucleic acids/proteins are transferred onto the membrane and are immobilized on the support. The probes are bound to target molecules on the support and various techniques are used for visualization (Hayes et al., 1989).

8.2 SOUTHERN BLOTTING

This technique of blotting was established by Dr. Edwin Southern. It is widely used in molecular biology for identification of target sequence of DNA from the DNA sample. For this, first the DNA electrophoresis is carried out then hybridization is performed in blotting membrane after which there is transfer step in which transfer of DNA takes place from gel to blotting membrane (Southern, 2006).

8.2.1 HISTORICAL CONTEXT

DNA strand has the ability to hybridize with its complementary structure as described in Watson Crick's model of DNA structure. Various experiments have been performed to understand renaturation ability of DNA and form a DNA/RNA hybrid through which one can check cell or tissue specific expression. Development of probes was another milestone which can identify specific sequences inside the target. Then target needed to get immobilize on solid support, like nitrocellulose powder, resulted in formation of nitrocellulose membrane. Southern then demonstrated that separated DNA fragments from gel electrophoresis, transferred to a blotting membrane, characterize DNA through their size and hybridization pattern. Electrophoresis can be done using polyacrylamide gel (PAGE) containing urea which can result in the transfer of the same amount of DNA to blotting paper. Sodium dodecyl sulfate (SDS) containing urea can also be used for electrophoresis with the advantage that resolution of bandwidth remains same so it can even identify 100 pg of DNA. To summarize southern blotting double-stranded DNA is formed containing a strand from target DNA and other strand from DNA probe produced *in vitro* (Brown, 1993).

8.2.2 PRINCIPLE

DNA is broken into small fragments using enzyme restriction endonucleases (ER). Gel electrophoresis separates these fragments made by ER according to their size. After this, transfer of DNA fragments to the blotting paper takes place and there it is incubated with highly selective probes. Probes can bind with target fragment with a resolution of one in a million.

8.2.3 PROCEDURE

The whole process of southern blotting is divided into the following steps:

1. **DNA Purification:** Cell lysis is done to extract DNA present inside the nucleus. It is done by incubating cell culture in detergent, resulting in a mixture containing DNA, protein, and debris. Proteinase enzyme lysis proteins present in mixture then DNA is purified and separated by alcohol precipitation. Besides manual isolation, commercially available kits can be used for the isolation of DNA from different sources (mammalian cells, plants cells, bacteria, and fungi) with the advantage of high purity.
2. **Fragmentation:** The long nucleotide sequences should be broken into smaller fragments, which is done by the restriction endonuclease enzyme. Partial digestion leads to inaccurate analysis so enzyme should be added in surplus amount, but enzyme concentration should not surpass the 1/10th of the total volume of digest. Master mixes should be made to avoid pipetting errors in case of many samples. The digests are incubated at 37°C in a water bath for 1–2 hours (for DNA samples obtained from cloning) or overnight in case of genomic DNA. After digestion traces of ethanol should be removed and the concentration of DNA is checked.
3. **Gel Electrophoresis:** Nucleic acids move toward the anode and their movement of the DNA fragments results in their separation by size. Around 0.7–2% gel is mostly used and bubble formation should be avoided.
4. **Denaturation:** Double stranded DNA is obtained after electrophoresis but we need single stranded DNA for hybridization. Alkaline solution is used for the formation of single stranded denatured DNA.
5. **Blotting:** Transfer of fragmented DNA to nitrocellulose known as blotting can be done through capillary or electro blotting. Nitrocellulose membrane were first used for southern blot, but nylon membranes are in use recently because of their ability to attach more quantity of DNA efficiently which allow it to bind with much less target DNA. NaCl solution is used for saturating DNA molecule after which permanent fixation is done through drying or UV radiation. This process for preparation takes 2 hours approximately. Then gel is placed under UV light at 256 nm to

check DNA and if separation does not take place gel should run for a long time. After that complete denaturation of gel takes place by submerging it in 1.5 M NaCl and 0.5 M NaOH for 15–30 minutes in a rocking shaker. After that this solution is replaced with 3 M NaCl and 0.5 M Tris HCl (pH 7) and incubated for 15–30 min resulting in neutralization of the gel. A thick filter dipped in 20X saline-sodium citrate (SSC) buffer (3.0 M NaCl and 0.3 M sodium citrate) is used with nitrocellulose filter, equal to the length of the strip of the gel, dipped in 2X SSC is used for transfer. Large filter paper is placed on plastic platform previously soaked in 20X SSC and glass sheet is placed on it. Gel is removed afterward neutralizing solution and is placed parallel to glass 2–3 mm apart approximately and opposite side of gel, second glass placed 2–3 mm apart nitrocellulose membrane is placed on top portion of gel with its ends balanced on glass. Finally absorbent paper is laid on top. Transfer takes approximately three hours. 20X SSC should be replenished to prevent drying during transfer and after transfer gel and nitrocellulose membrane should be attached. Boundaries of gel on nitrocellulose membrane are traced with pencil and gel is removed afterward if some piece remain attached it is removed with blade. The portion thus obtained is dissolved in 2X SSC for 20 minutes before baking in a vacuum oven for 2 hours at 80°C. After transfer hybridization takes place but for that there is pre-hybridization which is blocking non-specific reactions. This pre-hybridization solution has salmon sperms DNA mostly which act as a blocking agent in the hybridization chamber (Stachelin & Gordon, 1979).

6. Hybridization: In membrane buffer, labeled probe is added and incubated for 1 to 16 hours so that probe can find target sequence. As radioactive RNA is isolated from cells so it is present in small amounts. Labeling method depends on the source form which probe is derived could be of DNA (obtained from PCR product or cloning) or RNA (from cells). Membrane is immersed in hybridization solution [0.5% SDS, 6X SSC, 5X Denhardt's solution (for 1 X Denhardt's solution: 0.02% Ficoll, 0.02% polyvinylpyrrolidone and 0.02% bovine serum albumin (BSA)) and 100 mg/ml sheared, denatured salmon sperm DNA or yeast tRNA] with paraffin oil at 40–65°C (depending on the solvent). After which solvent is

removed and the probe is added and left overnight, the amount of which depends on membrane size. After hybridization remove the membranes, blotted against filter paper, washed with solvent for half hour causing removal of nonspecific hybrids. RNase treatment is used in case of background and membrane is air dried finally. Hybridization of florescent probe is visualized by autoradiography on an X-ray film and incase of chromogenic detection formation of color on membrane is checked. Detection time is 1 to 48 hours, depending on probe (Brown, 2001).

7. Reuse of Blots: For clinical samples with little target DNA blots can be reused. Warm alkali treatment can remove the probe, then the membrane is neutralized and hybridized with different probe.

8.2.4 SOUTHERN BLOTTING IN COMPARISON WITH PCR AND CHROMATOGRAPHY

PCR is dependent on information about DNA sequence for designing primers which is not required in Southern blotting. Southern blot method amplifies the level of detection limit necessary for searching specific locus in the genome in comparison to chromatography. But needs starting target DNA in high concentration so its drawback of southern blotting (Southern, 1975).

8.2.5 APPLICATIONS

Southern blotting is primarily used to search target DNA in DNA sample. It can also be used to identify certain viral infections, bacterial infections, isolate specific DNA, gene mutation, gene rearrangement, diagnosis neonatal and genetic disease. It can also be used in paternity analysis, phylogenetic studies, personal identification, and forensic studies.

Southern blotting helps to know the structure of particular gene, elucidation of restriction enzyme maps, finding methylated sites in genes. Application of restriction nucleases (like MspI and HpaII) can identify as well as cleave in identical sequence and thus helpful in mapping genomes needed to be mapped, also it can tell about clonal rearrangements

in immunoglobulin and T Cell receptor genes so is helpful to tell about immune response.

DNA fragments transfer is fast but efficiency in southern blotting is low because of quick dehydration of agarose gel as fluid is removed or withdrawn from both sides. It works finest when target sequence exists in high concentration (cloned DNAs of plasmids, cosmids, bacteriophages, BACs, and PACs) and less complex genome (*Drosophila* and *S. cerevisiae*) (Green & Sambrook, 2021) (Figure 8.1).

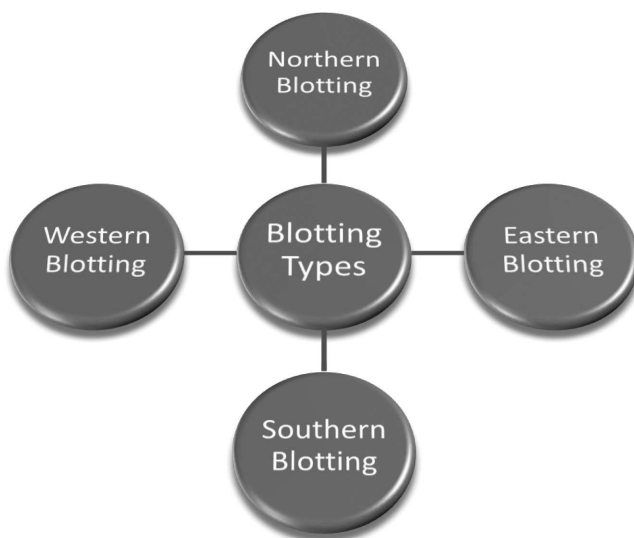


FIGURE 8.1 The basic types of blotting – Southern, Northern, Eastern, and Western blotting.

8.3 NORTHERN BLOTTING

Northern blotting is a technique employed widely to get insight in the abundance and size of specific RNA. The fractionation of RNA is carried out by gel electrophoresis and later on it is transferred onto the membrane where it is subjected to further analysis through labeled probes which are highly specific for particular RNA. Northern blotting (Alwine et al., 1977, 1979) is a further extension of Southern blotting, a technique developed by Ed Southern (Southern, 1975) for analysis of DNA. So, its name has no logical background as its name does not specify the name of its founder.

Northern blotting gives information about the length of molecules of RNA and also the existence of variants of different length because the electrophoresis of RNA takes place under denaturing conditions. Special conditions and procedures are required for circular or branched RNA molecules as they have aberrant mobility (Josefsen & Nielsen, 2011). Northern blotting is often done to detect the presence of specific RNA but is also gives quantitative information. RNA decay is measured through the addition of transcription inhibitors at various times. Northern blotting, as a quantitative method is used to compare the level of RNA rather than absolute quantification in different experimental conditions.

The RNA can be of any kind but in most of the cases total RNA extracted by guanidinium thiocyanate (Chomczynski & Sacchi, 1987) method or purification is done through affinity matrix for RNA preparation. Oligo (dT) chromatography is used to isolate polyA⁺ mRNA when sensitivity of RNA under consideration is an issue (Aviv & Leder, 1972), and then fractionation is done by gel electrophoresis. Cytoplasmic RNA is 3% of total RNA in the human cells. Relevant RNA can be isolated through this fraction and can be loaded onto the gel and RNA which is of less abundance can also be detected through this method. Simple protocols of fractionation which are based upon binding of RNA with column enhances the gels capacity and can be used effectively for microRNAs.

Agarose gel is used for fractionation of large size RNAs and with range of agarose concentration between 0.8%–1.4%. Denaturing conditions are used for this process because RNA being single stranded can fold into complex structures and becomes double stranded at some points too. The RNA fractionation thus depends upon both folding of RNA and its length. Glyoxal (McMaster & Carmichael, 1977) and formamide (Lehrach et al., 1977) are used commonly for denaturation of agarose gels. Formaldehyde is used in the gel and loading buffer as RNA denaturation is done by formamide and denatured state is maintained by formaldehyde as it makes covalent bonds with amine groups of guanine, cytosine, and adenine bases. Bases are involved in the base pair formation, so bonding with them prevents the secondary structure formation. Filter treatment is done after blotting to reverse the reaction so base pairing can occur with the probe. Glyoxal can be used only for pre-treatment of samples. It also causes denaturation by base pairing with guanine and its effect is also reversed by filter treatment after blotting. Sharper bands are obtained with glyoxal as compared to formaldehyde (Reijnders et al., 1973).

Northern blotting is often used to compare the level of RNA in different experimental conditions and each condition is represented by lane on the gel. Equal sample loading is of critical importance and on the basis of spectrophotometer equal amounts are used in case of whole cellular RNA. This works effectively as cytoplasmic RNA constitutes up to 80% of total RNA and experimental variables are unlikely to affect it. Northern blotting is much sensitive to degradation of sample so visual inspection of RNA is of crucial importance after electrophoresis. RNA like DNA is also stained by ethidium bromide, but staining is less efficient because intercalation does not take place and bonding occurs with ionic interaction. Ethidium bromide can be added into loading buffer in case of formaldehyde gels as inspection can be done during gel running. This cannot be done with glyoxal gels as there is a reaction between glyoxal and ethidium bromide and RNA cannot be denatured. It is not recommended to include ethidium bromide in the gel as there is migration of ethidium bromide in the gel and a stain is created in front of the gel.

8.3.1 NORTHERN BLOTTING APPROACHES

There are two main approaches of northern blotting, electroblotting, and capillary blotting. The passive method of capillary transfer is just similar to Southern blotting (Southern, 1975) and it is much easy to perform as materials are readily available in every laboratory. High salt transfer buffer is drawn from the reservoir from the gel through a stack of papers or by some other absorbent material, and on the top of gel, the binding membrane of RNA is placed. RNA is carried out of the gel with the flow of buffer and then it gets trapped on the membrane. Few hours are required for this transfer and is usually done over-night.

Downward capillary transfer is also the simplest method (Chomczynski & Mackey, 1994) in which only 1 hour is required for transfer, but it can be left for a few hours as it does not cause any problem. Application of pressure above gel or vacuum just below the gel can be used to increase the rate of transfer. Some specialized equipment is also available commercially. Capillary transfer works efficiently with agarose gel, whereas in case of polyacrylamide gels the best method is electroblotting (Bittner, Kupferer, & Morris, 1980). The membrane and gel are placed into a cassette which is placed into the electrophoresis cell. RNA gets transferred onto the

membrane when there is the supply of electrical field just perpendicular to the cassette.

Semi-dry type is one of its variations in which membrane and gel are just sandwiched between filter paper which is damped with transfer buffer, and in some cases the process of electroblotting gets accomplished in less than an hour.

8.3.2 CHOICE OF FILTER MEMBRANE

Filter membrane type is of crucial importance in northern blotting. Different options are available but in most of the cases nylon membranes which are positively charged are used. Amino groups impart a positive charge to the surface of the membrane and thereby increasing the affinity of the membrane for nucleic acids and immobilizing them on the surface of the membrane. Special care must be taken while dealing with nylon membranes as they also give non-specific results due to non-specific binding with different probes. High binding capacity and tensile strength are the main features of nylon membranes and they are available with various pore sizes. Membranes which are charged on both sides are highly recommended.

Northern blotting is used for quantitative and qualitative analysis of RNA and in some cases, it is able to detect as little as 10,000 molecules and it is approximately equivalent to less than 100 ng of cellular RNA. It is laborious to perform experiments of nuclease protection and special care and optimization is required. Northern blotting is one of the important techniques which is quite easy to perform, simple to handle and cost and time effective.

8.3.3 CONSIDERATIONS

Special care must be taken with northern blotting because RNA is very sensitive to degradation. Gloves must be worn to prevent degradation from RNases and equipment should be sterilized. Equipment should be reserved specially for RNA to prevent any contamination. The electrophoresis tank should be cleaned properly with detergent and then must be dried with ethanol.

8.4 WESTERN BLOTTING

Western Blotting or Immunoblotting or protein blotting evolved from Northern blotting (Alwine et al., 1975); and Southern blotting (Southern, 1975). Towbin and his co-workers (1979) coined this name to maintain the tradition of geographic naming initiated by Southern (1975). In western blotting proteins are transferred from SDS polyacrylamide gel to nylon membrane and they are exact replica of proteins on gel, thus, provide base for a number of experiments (Towbin et al., 1979). The breakthrough in immunology took place with western blotting as antibodies can bind with proteins on the membrane.

The ability of resolution associated with gel electrophoresis achieved its full potential with western blotting as before western blotting the use of molecular probes against various proteins posed various challenges. The transfer of proteins on membrane and then detection of various antigens and proteins finds its numerous applications in biochemistry and life sciences. This technique (Laemmli, 1970) provides an opportunity to transfer and characterize also those proteins which are less abundant. Certain advantages which are associated with western blotting are: (i) the immobility of proteins on membrane and different ligands can access them; (ii) it is easy to handle wet membranes; (iii) it is possible to make replica of gels; (iv) The transfer takes place with only small number of reagents; (v) patterns are stored for a long time; and (vi) it is possible to carry out successive analysis with same transfer (Kost et al., 1994; Gershoni and Palade, 1982; Gershoni, 1988).

8.4.1 EFFICIENCY OF BLOTTING

Several factors such as gel natures, membrane being used and molecular mass of protein which is to be transferred effect the transfer of gel to the membrane. The use of thin gels such as those of acrylamide gives more efficient and robust results, but the problem arises with the use of ultra-thin gels as they are quite difficult to handle (Harlow & Lane, 1988). Blotting does not occur properly if the resolution of proteins of high molecular weight is done on SDS. Special buffers, heat, and partial digestion of such proteins is done to ensure efficient transfer from gel to solid membrane (Renart et al., 1979; Elkon et al., 1984; Gibson, 1981; Bolt & Mahoney, 1997; Kurien & Scofield, 2002).

8.4.2 EFFICIENCY OF BINDING WITH MEMBRANE

PVDF, nitrocellulose, activated nylon and activated paper are commonly used (Towbin et al., 1979; Elkon et al., 1984; Gershoni & Palade, 1983). Nitrocellulose is used commonly but proteins do not make covalent bonds with nitrocellulose membrane and it becomes brittle once dried. A small percentage of the actual amount of proteins bind with membrane as most of proteins with small molecular weight pass through the membrane. This problem can be tackled by using membrane having small pores (Suzuki et al., 2008). A small amount of polyester has been incorporated into supported nitrocellulose as it provides mechanical strength and makes it easy to deal with membrane.

Proteins make covalent bonds with activated paper, but the coupling method is not compatible with most of system of gel electrophoresis. In activated papers, linking takes place through primary amines and buffers must be free of any amino group in case of this paper. The paper is not cost effective and most of the reactive groups present on the paper have limited half-life. Nylon has great mechanical strength, but only a fraction of proteins can bind with membrane. Activated nylon can be used to overcome these issues, but it also causes non-specific binding.

Physical strength, high capacity of binding and chemical stability are main advantages associated with PVDF membranes. Proteins electroblotted on PVDF can make replicate lanes and can be used for various experimental purposes like sequencing of N-terminal and separation of internal sequences. PVDF membrane in contrast to nitrocellulose membrane can also be stained with Coomassie blue hence making it possible to excise proteins from N-terminal.

8.4.3 TECHNIQUES OF TRANSFERRING PROTEINS

Three ways are in use to transfer proteins from SDS-PAGE to PVDF or nitrocellulose membrane and they are western blotting (Towbin et al., 1979), vacuum assisted flow (Peferoen et al., 1982) and simple diffusion (Kurien & Scofield, 1997; Renart et al., 1979). A brief overview of these methods has been described as in subsections.

8.4.3.1 VACUUM BLOTTING

Vacuum blotting was used (Peferoen et al., 1982) as alternative to electroblotting or simple diffusion. Polypeptides and proteins on the gel were

transferred from gel to membrane through a suction pump attached to the dryer system of gel. Transfer of high and low molecular proteins can be achieved through this method. Nitrocellulose membrane with small size of pore is preferable for proteins with low molecular weight as these proteins are not adsorbed properly on membrane with large pore size.

Gel should be kept hydrated enough to avoid its stickiness with the membrane. There should be adequate buffer when time of transfer spans for more than 45 minutes. The concentration of gel is also of much importance in such cases.

8.4.3.2 DIFFUSION BLOTTING

Diffusion is simplest method which was developed to transfer proteins separated by iso-electric focusing on membranes (Kurien & Scofield, 1997; Reinhart & Malamud, 1982; Jägersten et al., 1988; De et al., 1990; Heukeshoven & Dernick, 1995; Olsen & Wiker, 1998; Chen & Chang, 2001, pp. 24–30). On the top of the gel, a membrane is placed and it is piled with stack of filter papers. Diffusion process is facilitated by glass plate and other object with specific weight placed on the top but it could not get popular as proteins were not transferred quantitatively. It gained a bit of importance when it was demonstrated that up to 12 blots can be obtained if the gel is sandwiched sequentially between two membranes (Kurien & Scofield, 1997). These membranes which are lifted from SDS PAGE gels are much effective for identification of proteins through mass spectroscopy (Gravel, 2008). Coomassie blue can be used to stain the gel after immunoblotting. Many imprints can be obtained from the same gel which makes it effective in comparison with electroblotting and identical imprints can be tested with varying antisera.

^{14}C labeled proteins are used to get quantitative information about proteins that were transferred by diffusion blotting. Around 10% transfer took place through 3 minutes of diffusion blotting in comparison with electroblotting. Shrinkage of gels takes place when they are subjected to diffusion blotting multiple times. Plastic support can help to address this issue (Chen & Chang, 2001; Olsen & Wiker, 1998).

Zymography or activity gel electrophoresis was studied with regard to diffusion, and in this case, enzymes are electrophoresed through polyacrylamide gel which is discontinuous and contains substrate of the enzyme.

SDS is removed after electrophoresis from the gel, and Triton X-100 is used for washing purposes. This causes degradation of substrate and renaturation of enzyme. Enzyme activity is detected after the gel is stained with CBB and clear lysis bands shown against a background of blue color (Bischoff et al., 1998). An additional analysis which uses another gel is required for examination of specific band, and this has been circumvented through diffusion blotting (Chen & Chang, 2001). Blotting of activity gel was done on PVDF for the purpose of immunostaining and after blotting, the gel was removed and the remaining gel can be used for other purposes.

Around 25–50% of (Chen & Chang, 2001) proteins are transferred through diffusion blotting as compared to that of electroblotting. Diffusion blotting is associated with significant loss of proteins during transfer. Compression and stretching are prevented as gel remains on plastic which acts as support which ensures reliable determination of molecular mass. Diffusion blotting is adopted when quantitative transfer of proteins is not required.

8.4.4 ELECTROBLOTTING

Electroblotting is commonly used for transfer of proteins from gel to membrane because of efficiency and speed associated with it. The electro-elution can be done by placing the sandwich of membrane and gel between absorbent papers which are soaked in the transfer buffer or by immersion of sandwich of gel-membrane completely in buffer.

8.4.5 SEMI-DRY TRANSFER

The sandwich of gel-membrane is placed between electrodes of carbon plate in this method. This is also called horizontal blotting and it uses two plate electrodes to ensure uniform electrical field over the short distance. Sandwiches of membrane-gel are soaked properly in the transfer tube. The main pros which are associated with this assembly are: (i) there can be cheap blocks of carbon as electrodes; (ii) there is no requirement of high power; and (iii) it is possible to simultaneous blotting of gels (Kurien & Scofield, 2006).

8.4.6 INTEGRATED CHIPS FOR WESTERN BLOTTING

The microchip has been developed for the integration of separation and blotting into single microfluidic platform. A large number of processes are integrated into a single chip through a microfluidic analysis platform and it offers comparative advantages (Ríos et al., 2012; Patabadige et al., 2016). These processes are much efficient and faster as compared to other macroscopic techniques and there is substantial reduction in requirement of various reagents. Significant work has been done by Herr and his coworkers for miniaturization of protein analysis onto the microchips (He & Herr, 2009, 2010; Chen et al., 2011; He et al., 2011; Tia et al., 2011; Hughes & Herr, 2012; Hughes et al., 2012; Chen et al., 2013; Kim et al., 2012; Duncombe & Herr, 2013; Herr, 2013; Tentori et al., 2013; Gerver & Herr, 2014; Kang et al., 2014).

Separation and blotting (He & Herr, 2009, 2010; Tia et al., 2011; Hughes et al., 2012; Kim et al., 2012; Chung et al., 2013) were integrated initially into microchips with various reservoirs were connected to glass microfluidic chip for introduction of buffer, sample, and voltage programs (Hughes et al., 2012). There is thin PA-gel housed in the central chamber and it is modified for creation of zones for loading sample, separation, and blotting. The functionality and pore size of the gel are modified by multi-step photopolymerization. Early work was designated for immobilization of probes antibodies in the region of blotting (He & Herr, 2009, 2010).

The samples of native proteins are introduced at the loading zone, and their separation is done on the basis of charge-to-mass ratio. Proteins after separation are transferred laterally onto the blotting zone through applied voltage. The blotting region contains immobilized probes and they capture the protein of interest while others move away (Sanders et al., 2016).

8.5 EASTERN BLOTTING

Biochemical technique employed for identification of post-translational modifications (PTMs) in proteins (addition of phosphates, lipids, and glycoconjugates). It was developed in 1979 by Towbin. It is an extension of western blot used mostly for epitope detection in carbohydrates.

8.5.1 PRINCIPLE

Gel electrophoresis is used to separate protein samples, transferred to nitrocellulose membrane, primary antibodies are added which bind nitrocellulose membrane, removed unbounded primary antibody through washing the membrane, added enzyme-labeled secondary antibody, and finally binding with primary antibody.

There are many definitions of eastern blotting and all of them are derivatives of western blot developed by Towbin in 1979.

8.5.2 MIDDLE-EASTERN BLOTTING

It is a blot of poly-A RNA resolved by agarose. RNA is then immobilized and probed using DNA.

8.5.3 EASTERN-WESTERN BLOTTING

It is blotting phospholipids on PVDF (polyvinylidene fluoride) or nitrocellulose membrane before proteins transfer to nitrocellulose membrane through western blotting after which it is probed specific antibodies for confirmation.

8.5.4 FAR-EASTERN BLOTTING

It is based on staining lipids with lectin or antibody, transferred to PVDF membranes. Eastern blotting can be used to detect glycans just like lectins are used to detect protein glycosylation also known as lectin blotting. Glycoconjugates detection by blotting BSA to PVDF membrane, which is followed by treatment with periodate. Protein is thus oxidized then it is reacted with some complex mixture thus resulting in generation novel conjugate on membrane. This membrane is probed with antibodies for finding epitopes of interest.

8.5.4.1 PROCEDURE

The whole process is divided in following steps:

- i. **Separation:** In the first step, gel electrophoresis is used to separate the test sample.
- ii. **Transfer:** In the next step transferred molecules horizontally to nitrocellulose membrane.
- iii. **Addition of Primary Antibody:** Primary antibody is added to solution after transfer is completed.
- iv. **Washing:** The membrane is washed for removing primary antibody which was not bound.
- v. **Addition of Secondary antibody:** Enzyme-labeled secondary antibody is added, emitting fluorescence once product is formed.
- vi. **Autoradiography:** The molecule of interest is confirmed by performing autoradiography.

8.5.5 APPLICATIONS

As proteins undergo modification before becoming functional, these modifications are called post-translational modifications (PTMs). These modifications include biotinylation, sulfinylation, alkylation, formylation, S-nitrosylation, acetylation, glutamylation, acylation, isoprenylation, arginylation, ADP-ribosylation, glycosylation, geranylgeranylation, succinylation, glycosylation, hydroxylation, methylation, lipoylation, phosphorylation, nitroalkylation, selection, ubiquitination, phosphopantetheinylation, transglutamination, prenylation, sulfation, and sulfonylation. N terminus PTMs play part in translocation across the sides of biological membranes. These proteins could be secretory or incorporated in cell or organelle membranes (lysosomes, mitochondria chloroplast and plasma membrane). Post translated proteins expression is altered in many diseases. Eastern blotting can identify protein modifications in bacterial species. Non-virulent *E. Muris* antigenic proteins had more post translational modifications than highly virulent IOE. Concanavalin A (bind glycans-containing mannose), Cholera toxin B subunit (binds gangliosides), and nitrophosphomolybdate-methyl green (phosphoproteins binding) can be utilized for identifying protein modifications.

KEYWORDS

- **bovine serum albumin**
- **polyacrylamide gel**
- **polyvinylidene fluoride**
- **post-translational modifications**
- **saline-sodium citrate**
- **sodium dodecyl sulfate**

REFERENCES

- Alwine, J. C., Kemp, D. J., & Stark, G. R., (1977). Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proceedings of the National Academy of Sciences*, 74(12), 5350–5354.
- Alwine, J. C., Kemp, D. J., Parker, B. A., Reiser, J., Renart, J., Stark, G. R., & Wahl, G. M., (1979). [15] Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzyloxymethyl paper. In: *Methods in Enzymology* (Vol. 68, pp. 220–242). Academic Press.
- Aviv, H., & Leder, P., (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proceedings of the National Academy of Sciences*, 69(6), 1408–1412.
- Bischoff, K. M., Shi, L., & Kennelly, P. J., (1998). The detection of enzyme activity following sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Analytical Biochemistry*, 260(1), 1–17.
- Bittner, M., Kupferer, P., & Morris, C. F., (1980). Electrophoretic transfer of proteins and nucleic acids from slab gels to diazobenzyloxymethyl cellulose or nitrocellulose sheets. *Analytical Biochemistry*, 102(2), 459–471.
- Bolt, M. W., & Mahoney, P. A., (1997). High-efficiency blotting of proteins of diverse sizes following sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Analytical Biochemistry*, 247(2), 185–192.
- Brown, T. A., (2001). Southern blotting and related DNA detection techniques. *e LS*.
- Brown, T., (1993). Southern blotting. *Current Protocols in Molecular Biology*, 21(1), 2–9.
- Chen, H., & Chang, G. D., (2001). Simultaneous immunoblotting analysis with activity gel electrophoresis in a single polyacrylamide gel. *Electrophoresis*, 22(10), 1894–1899.
- Chen, J. Q., Heldman, M. R., Herrmann, M. A., Kedei, N., Woo, W., Blumberg, P. M., & Goldsmith, P. K., (2013). Absolute quantitation of endogenous proteins with precision and accuracy using a capillary Western system. *Analytical Biochemistry*, 442(1), 97–103.
- Chen, X., Kapil, M. A., Hughes, A. J., & Herr, A. E., (2011). Single-microchannel, multistep assay reports protein size and immunoaffinity. *Analytical Chemistry*, 83(17), 6573–6579.

- Chomczynski, P., & Mackey, K., (1994). One-hour downward capillary blotting of RNA at neutral pH. *Analytical Biochemistry*, 221(2), 303–305.
- Chomczynski, P., & Sacchi, N., (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry*, 162(1), 156–159.
- Chung, M., Kim, D., & Herr, A. E., (2013). Microchamber Western blotting using poly-L-lysine conjugated polyacrylamide gel for blotting of sodium dodecyl sulfate coated proteins. *Analytical Chemistry*, 85(16), 7753–7761.
- De Keyser, F., Verbruggen, G., Veys, E. M., Nimmegeers, J., Schatteman, L., Goethals, K., & Vandenbossche, M., (1990). “Microgel diffusion blotting” for sensitive detection of antibodies to extractable nuclear antigens. *Clinical Chemistry*, 36(2), 337–339.
- Duncombe, T. A., & Herr, A. E., (2013). Photopatterned free-standing polyacrylamide gels for microfluidic protein electrophoresis. *Lab on a Chip*, 13(11), 2115–2123.
- Elkon, K. B., Jankowski, P. W., & Chu, J. L., (1984). Blotting intact immunoglobulins and other high-molecular-weight proteins after composite agarose-polyacrylamide gel electrophoresis. *Analytical Biochemistry*, 140(1), 208–213.
- Gershoni, J. M., & Palade, G. E., (1982). Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to a positively charged membrane filter. *Analytical Biochemistry*, 124(2), 396–405.
- Gershoni, J. M., & Palade, G. E., (1983). Protein blotting: Principles and applications. *Analytical Biochemistry*, 131(1), 1–15.
- Gershoni, J. M., (1988). Protein blotting: A manual. *Methods of Biochemical Analysis*, 33, 1–58.
- Gerver, R. E., & Herr, A. E., (2014). Microfluidic western blotting of low-molecular-mass proteins. *Analytical Chemistry*, 86(21), 10625–10632.
- Gibson, W., (1981). Protease-facilitated transfer of high-molecular-weight proteins during electrotransfer to nitrocellulose. *Analytical Biochemistry*, 118(1), 1–3.
- Gravel, P., (2008). Protein blotting. In: *Molecular Biomethods Handbook* (pp. 365–375). Humana Press.
- Green, M. R., & Sambrook, J., (2021). Southern blotting. *Cold Spring Harbor Protocols*, 2021(7), pdb-prot100487.
- Harlow, E. D., & Lane, D., (1988). *A Laboratory Manual* (p. 579). New York: Cold Spring Harbor Laboratory.
- Hayes, P. C., Wolf, C. R., & Hayes, J. D., (1989). Blotting techniques for the study of DNA, RNA, and proteins. *BMJ: British Medical Journal*, 299(6705), 965.
- He, M., & Herr, A. E., (2009). Microfluidic polyacrylamide gel electrophoresis with in situ immunoblotting for native protein analysis. *Analytical Chemistry*, 81(19), 8177–8184.
- He, M., & Herr, A. E., (2010). Automated microfluidic protein immunoblotting. *Nature Protocols*, 5(11), 1844–1856.
- He, M., & Herr, A. E., (2010). Polyacrylamide gel photopatterning enables automated protein immunoblotting in a two-dimensional microdevice. *Journal of the American Chemical Society*, 132(8), 2512–2513.
- He, M., Novak, J., Julian, B. A., & Herr, A. E., (2011). Membrane-assisted online renaturation for automated microfluidic lectin blotting. *Journal of the American Chemical Society*, 133(49), 19610–19613.

- Herr, A. E., (2013). Disruptive by design: A perspective on engineering in analytical chemistry. *Analytical Chemistry*, 85(16), 7622–7628.
- Heukeshoven, J., & Dernick, R., (1995). Effective blotting of ultrathin polyacrylamide gels anchored to a solid matrix. *Electrophoresis*, 16(1), 748–756.
- Hughes, A. J., & Herr, A. E., (2012). Microfluidic western blotting. *Proceedings of the National Academy of Sciences*, 109(52), 21450–21455.
- Hughes, A. J., Lin, R. K., Peehl, D. M., & Herr, A. E., (2012). Microfluidic integration for automated targeted proteomic assays. *Proceedings of the National Academy of Sciences*, 109(16), 5972–5977.
- Jägersten, C., Edström, A., Olsson, B., & Jacobson, G., (1988). Blotting from PhastGel media after horizontal sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Electrophoresis*, 9(10), 662–665.
- Josefsen, K., & Nielsen, H., (2011). Northern blotting analysis. In: *RNA* (pp. 87–105). Humana Press.
- Kang, C. C., Lin, J. M. G., Xu, Z., Kumar, S., & Herr, A. E., (2014). Single-cell western blotting after whole-cell imaging to assess cancer chemotherapeutic response. *Analytical Chemistry*, 86(20), 10429–10436.
- Kim, D., Karns, K., Tia, S. Q., He, M., & Herr, A. E., (2012). Electrostatic protein immobilization using charged polyacrylamide gels and cationic detergent microfluidic western blotting. *Analytical Chemistry*, 84(5), 2533–2540.
- Kost, J., Liu, L. S., Ferreira, J., & Langer, R., (1994). Enhanced protein blotting from PhastGel media to membranes by irradiation of low-intensity ultrasound. *Analytical Biochemistry*, 216(1), 27–32.
- Kurien, B. T., & Scofield, R. H., (1997). Multiple immunoblots after non-electrophoretic bidirectional transfer of a single SDS–PAGE gel with multiple antigens. *Journal of Immunological Methods*, 205(1), 91–94.
- Kurien, B. T., & Scofield, R. H., (2002). Heat-mediated, ultra-rapid electrophoretic transfer of high and low molecular weight proteins to nitrocellulose membranes. *Journal of Immunological Methods*, 266(1, 2), 127–133.
- Kurien, B. T., & Scofield, R. H., (2006). Western blotting. *Methods*, 38(4), 283–293.
- Laemmli, U. K., (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259), 680–685.
- Lehrach, H., Diamond, D., Wozney, J. M., & Boedtker, H., (1977). RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. *Biochemistry*, 16(21), 4743–4751.
- McMaster, G. K., & Carmichael, G. G., (1977). Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. *Proceedings of the National Academy of Sciences*, 74(11), 4835–4838.
- Olsen, I., & Wiker, H. G., (1998). Diffusion blotting for rapid production of multiple identical imprints from sodium dodecyl sulfate-polyacrylamide gel electrophoresis on a solid support. *Journal of Immunological Methods*, 220(1, 2), 77–84.
- Patabadige, D. E., Jia, S., Sibbitts, J., Sadeghi, J., Sellens, K., & Culbertson, C. T., (2016). Micro total analysis systems: Fundamental advances and applications. *Analytical Chemistry*, 88(1), 320–338.

- Peferoen, M., Huybrechts, R., & De Loof, A., (1982). Vacuum-blotting: A new simple and efficient transfer of proteins from sodium dodecyl sulfate—Polyacrylamide gels to nitrocellulose. *FEBS Letters*, 145(2), 369–372.
- Reijnders, L., Sloof, P., Sival, J., & Borst, P., (1973). Gel electrophoresis of RNA under denaturing conditions. *Biochimica et Biophysica Acta (BBA)-Nucleic Acids and Protein Synthesis*, 324(3), 320–333.
- Reinhart, M. P., & Malamud, D., (1982). Protein transfer from isoelectric focusing gels: The native blot. *Analytical Biochemistry*, 123(2), 229–235.
- Renart, J., Reiser, J., & Stark, G. R., (1979). Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: A method for studying antibody specificity and antigen structure. *Proceedings of the National Academy of Sciences*, 76(7), 3116–3120.
- Rios, Á., Zougagh, M., & Avila, M., (2012). Miniaturization through lab-on-a-chip: Utopia or reality for routine laboratories? A review. *Analytica Chimica Acta*, 740, 1–11.
- Sanders, B. J., Kim, D. C., & Dunn, R. C., (2016). Recent advances in microscale western blotting. *Analytical Methods*, 8(39), 7002–7013.
- Southern, E. M., (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.*, 98(3), 503–517.
- Southern, E., (2006). Southern blotting. *Nature Protocols*, 1(2), 518–525.
- Staehelin, T., & Gordon, J., (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. USA*, 76, 4350–4354.
- Suzuki, Y., Takeda, Y., & Ikuta, T., (2008). Immunoblotting conditions for human hemoglobin chains. *Analytical Biochemistry*, 378(2), 218–220.
- Tanaka, H., Putalun, W., & Shoyama, Y., (2012). Fingerprinting of natural product by eastern blotting using monoclonal antibodies. *Chromatography Research International*, 2012.
- Tentori, A. M., Hughes, A. J., & Herr, A. E., (2013). Microchamber integration unifies distinct separation modes for two-dimensional electrophoresis. *Analytical Chemistry*, 85(9), 4538–4545.
- Tia, S. Q., He, M., Kim, D., & Herr, A. E., (2011). Multianalyte on-chip native western blotting. *Analytical Chemistry*, 83(9), 3581–3588.
- Towbin, H., Staehelin, T., & Gordon, J., (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proceedings of the National Academy of Sciences*, 76(9), 4350–4354.

CHAPTER 9

Chromosome Jumping

MUHAMMAD ISHTIAQ, MAHNOOR MUZAMMIL, and
MEHWISH MAQBOOL

*Department of Botany, Mirpur University of Science and Technology
(MUST), Mirpur, Azad Jammu and Kashmir (AJK), Pakistan*

ABSTRACT

The present chapter provides a brief description of the chromosome jumping technique. The technique was introduced by for the formation of genetic markers with gene clones with larger genes with hundreds of kilobases size. With the advent of respective technique formation of genetic marker for genetic disorders was made possible, which couldn't be achieved using other present techniques. Along with this chromosome jumping allows formation of jumping library with gene have greater base pair distance.

Development and advance in molecular cloning and DNA analysis has contributed major evolution in genomic analysis. From single gene with few kilobases size to larger genes with hundreds of kilobases size are sequenced and cloned with in either cosmid or plasmid or by using chromosome walking technique. On the other hand, while dealing with larger cluster cloning chromosome walking has beaten other techniques. Present conventional techniques don't allow separation of base pairs at greater (millions of base-pair) distance. Such greater distant jumping is provided by two innovated techniques, i.e., PFG gel electrophoresis and Chromosome Jumping. The present chapter will focus on 'Chromosome jumping.'

Chromosome jumping is a tentative approach in molecular biology used for analyzation of physical Genome project. Technique involves the linking and jumping of DNA fragments for the construction of specific restricted markers or genes. Technique uses 'jumping' and 'linking' of specific DNA fragments for the generation of genomic libraries. It is an advanced type of chromosome walking that allows breaking of the chromosomes at specific breakpoints. Chromosomes break as guided and permit by the cloning of DNA ends only, without the middle section involvement. This is achieved either by rearrangement of pre-existing DNA/genome or may involve partial digestion of DNA ends using restriction endonuclease. In some cases, enzymes are used for cutting of DNA ends.

Chromosome jumping allows jump over hundreds of kilobases as in this technique the jumping is independent of the vector's capacity. Technique is more suitable for larger DNA fragments with several kilobases in size. Along with this technique is also involved in gene mapping. Chromosome jumping is accomplished due to the presence of transposable elements TEs, these TEs are found in both prokaryotes and eukaryotes (Figure 9.1).

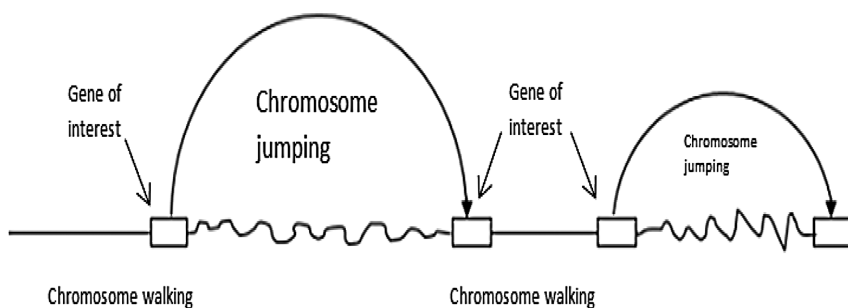


FIGURE 9.1 Describing mode of jumping mechanism of chromosomes.

9.1 ORIGIN OF JUMPING CHROMOSOMES

The concept of jumping chromosomes was introduced by an American scientist Barbara McClintock in around 1930s. Before this gene were considered as fixed entities, like beads on a string. McClintock, for the first time describes that some gene are capable of moving from their respective designated places, i.e., some genes have mobility. McClintock named these mobile genes as transposable elements TEs. TES were first

observed in the *Zea mays* (maize) model plant. *Zea mays* were selected as model plants as it produces a number of varying kernels. Each kernel behaves as an embryo and is produced via fertilization. So, plant produces a wide range of embryos at single ear hence making maize as an ideal model plant. Discovery of TEs also called jumping genes in maize was an innovative progress in genetics. This discovery give rise to the idea that gene present on chromosomes are not stationary but can alternate or rearrange themselves on DNA structure. Conversely, the idea of TEs was widely accepted and McClintock was bestowed with Nobel Prize in 1983 (Pray & Zhaurova, 2008; Sandeep, 2012).

9.2 GENERAL WORKING

Chromosome jumping is an extended version of “chromosome walking” that allows the DNA cuttings at longer distances. Chromosome jumping is performed on longer DNA fragments, i.e., of several Kilobase pairs kb (Slotkin & Martienssen, 2007). Generally, the technique involves following steps:

- First step in chromosome jumping involves the identification of TEs gene from desire genome.
- After this the special frequent-cutting Restriction enzymes (MboI or BamHI) are used for cutting the gene sequence.
- The resultant fragment size is carefully chosen to be around N kb in length (as chromosome jumping requires higher molecular weight molecules with several KB lengths).
- Then DNA is ligated at lower concentration, this results in the favored ligation in circular pattern rather in multimers form.
- Circulation of DNA fragment is an important step in chromosome jumping which is accomplished by using DNA ligase.
- At this stage DNA markers (amber suppressor tRNA gene supF) are added for the selection of the desired fragment from circular DNA molecule.
- Then another restriction enzyme (EcoRI) is added that causes the complete digestion of the larger selected fragment.
- After this the clones of DNA fragment are made in such a way that they contain vector generally *E. coli*. Vector, i.e., *E. coli* is

sequenced at the ligation sites (Poustka et al., 1987; Poustka & Lehrach, 1986).

- Vector should be selected in accordance to the DNA marker acquaints later.
- The Fragment that has been left is referred to as library of junction fragments or “jumping library.”
- Every clone in the jumping library has DNA sequence at one end, i.e., sequence at 2 and 3, of two adjacent cut end of enzyme Not I, while a clone in linking library has the DNA sequence present at either side of a single Not I site, e.g., regions 1 and 2 in clone 1 and regions 3 and 4 in clone 2.
- After this the library is screened via using probes which represents the “starting point” of desire “chromosome hop,” i.e., determination of the genome location for interrogation.
- Finally, this gives rise to the cloned fragments of DNA which is homologous to probe. These probes are generated with the help of DNA markers via, another DNA sequence that was originally located N kb away (thus being called “jumping”), as shown in Figure 9.1.
- By formation of different varying libraries of varying N value. This forms the entire genome for mapping allowing the movement of DNA fragment from one location to another, with complete control of the mobility (Figure 9.2).

This process allows the rapid mobility of desire gene cloning or genome cloning. This technique can be linked with chromosome walking for getting desire gene (Madireddy & Gerhardt, 2017).

9.3 DNA JUMPING LIBRARY

Chromosome jumping and chromosome walking are way far different techniques, although chromosome jumping is extension of chromosome walking. Chromosome library is not the same as chromosome walking, as in chromosome library DNA clones are managed before final instigation. In chromosome jumping library clones of DNA fragment are taken randomly from different locations irrespective of gap between selected genes from same chromosome (Collins et al., 1987; Drumm, 2001). Random fragments of DNA are ligated using a common adaptor in jumping library. Shorter fragments formed sometimes declaim difficulties

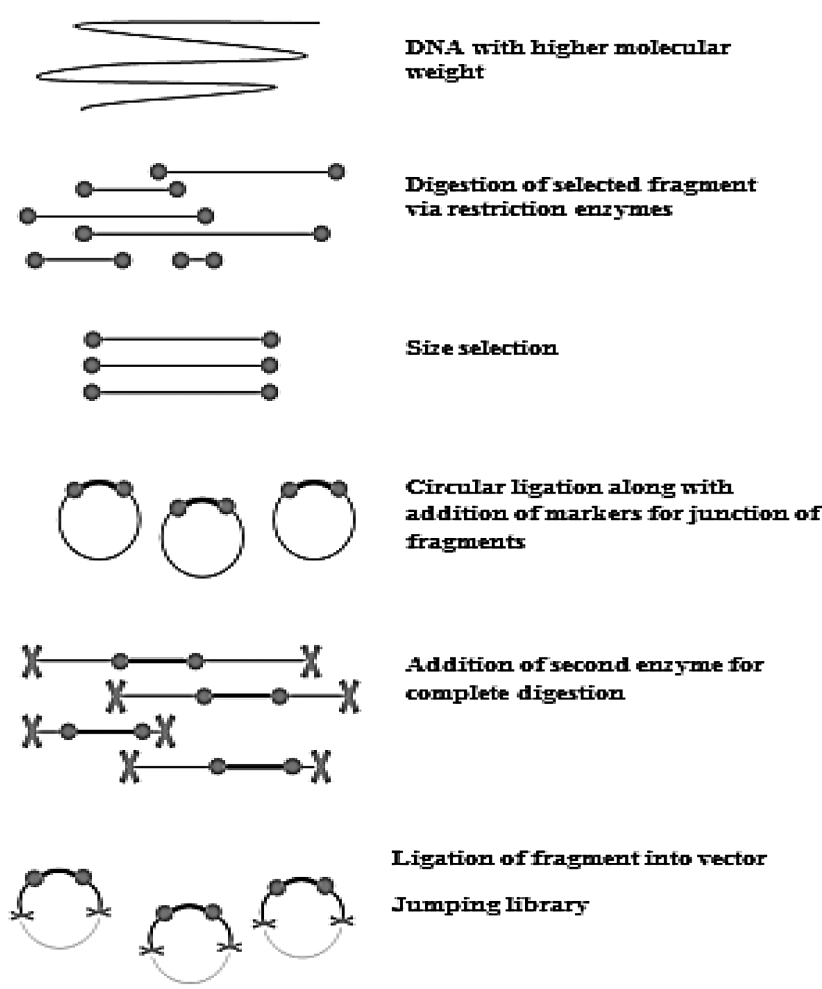


FIGURE 9.2 Mechanism of chromosome addition and deletion.

in structure identification of different deviates, like indels, translocations, and multiplication. While for larger DNA fragments, there is repetition of regions which causes more complications in alignment of fragments. In line of these issues jumping library is constructed using NGS that helps in the structural mapping of different variants and scaffold of de novo assemblages. The classification of jumping library is accomplished in accordance with the length of the DNA fragment (Figure 9.3).

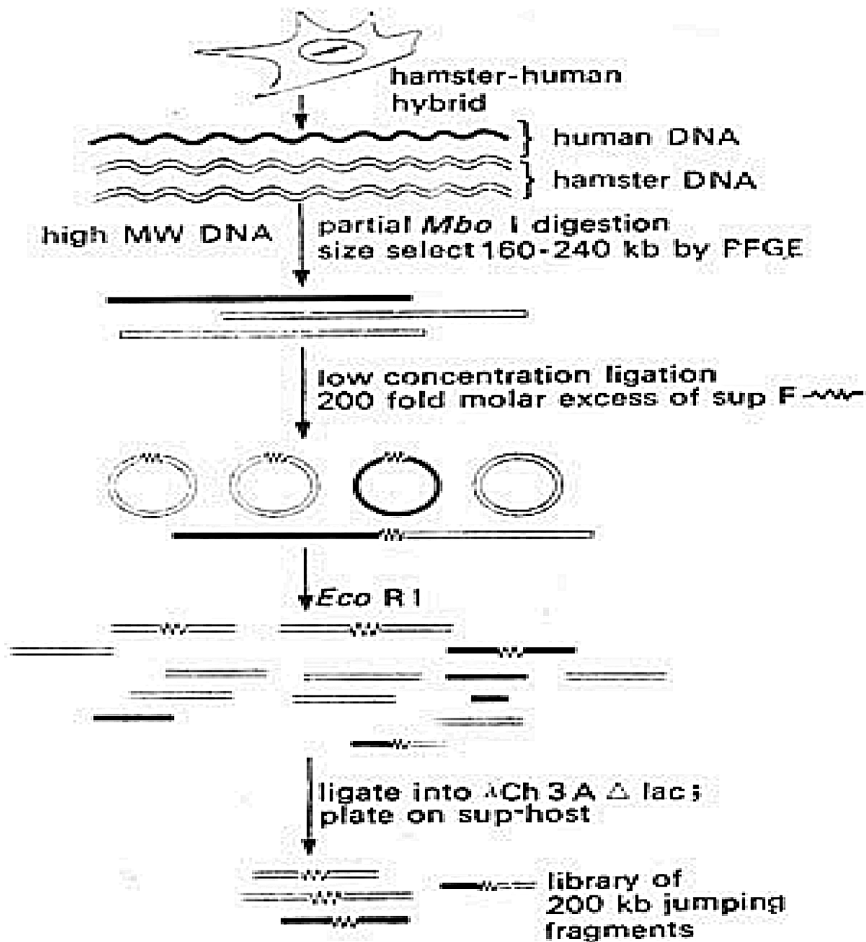


FIGURE 9.3 DNA jumping library shown, how genes move or jump.

9.4 APPLICATIONS

Before the advent of chromosome jumping technique, the sequencing of larger molecular clones was a great difficulty. Cloning and manipulation of smaller genetic marker was possible using chromosome walking, which helped a lot in the cloning of markers for different diseases. However, in 1987, after the successful discovery of chromosome jumping technique, this difficulty was overcome. Initially chromosome jumping technique

was implied for the cloning of genetic markers for cystic fibrosis disease that causes a recessive disease in autosomal chromosome. After the tremendous success of genetic marker for respective disease, the technique was adopted for different purposes and was proved as a useful technique for constructing jumping library/markers. Using chromosome jumping first marker for cystic fibrosis disease affecting chromosome number 7 was constructed. Cystic fibrosis gene causes down-streaming of 240 kb downstream of the met gene. Library screening of jumping clone started after the formation of this marker (Figure 9.3).

Another important aspect of chromosome jumping is that it decreases the mapping “steps” along with dodging of highly tedious regions of mammals’ genome. Along with this it also allows the formation of probes that can diagnose genetic diseases or disorders quickly (Figure 9.3).

Generally, the balanced chromosomal re-shuffling proved to be had a pertinent impact on diseases inception and occurrence ratio which has been demonstrated by the previous researchers work on leukemia and other infirmities. However, many of these reshuffling or rearrangement of genes on the chromosomes are undetected through chromosome micro-array technique. The other approaches used such as karyotype and FISH procedures can be useful to identify translocations and/or inversions on the chromosomes but these all processes are laborious, intensive time consuming and with low resolution results, because in this process many of small genomic alteration or changes are undetected or missed on image. Another new approach can be applied for the study of jumping library named as NGS combined form to locate and identify such genomic variations or changes. In a study, this method was used to determine the fine map as “a de novo” balanced process of translocation in a patient who was suffering from Wilms’ tumor infirmity (Talkowski et al., 2011). It was described by the authors that in the research ca. 50 million set of reads were produced but out of these only 11.6% replicates could be seen and mapped with reference genome library which is representation of a sixfold coverage of total analysis. Talkowski et al. in his research process compared various techniques to explore and find chromosomal alterations which showed that modification was present in “jumping library” in combination to next generation of DNA sequences which was proved an accurate procedure for chromosomal mapping through these breakpoints (Talkowski et al., 2011). There are two main varieties of jumping libraries, i.e., short-jump libraries and custom barcoded jumping libraries that are compared with

typical sequencing library. For standard NGS, 200–500 bp fragments are generated. Chimeric pairs are formed by 0.03%–0.54% fragments that are paired at end-reads which are mapped with two varying chromosomes. Consequently, precise DNA fragments cover the cutoff point area. With use of short jumping library with 3.2–3.8 kb fragment length, percentage of chimeric pairs enhances up to 1.3% (Sandeep, 2012).

9.5 PRENATAL DIAGNOSIS

Convention cytogenetic testing doesn't allow gene-level perseverance which is required for the prediction of the outcome of a pregnancy along with complete genome sequencing is not general practice for exploration of routine infirmity prenatal diagnosis. Whole-genome jumping library can be complementing, conservative for prenatal testing. This novel technique is proven successful in the treatment of charge syndrome.

9.6 GENETIC DISORDERS

Chromosome jumping library enables to overcome the addressing impediment for Primer walking that have higher molecular gaps in-between. Chromosome jumping is highly suitable for genetic disorder like Cystic fibrosis. Respected disorder is caused due to mutation on the 7th chromosome mutation. The detection of diseases on chromosome was difficult as it is composed of exon and non-coding DNA fragment in eukaryotic. Along with this, the major problem is the smaller size of introns much makes the detection more difficult. Here chromosome jumping allows the formation of the genetic markers with 100 kilobase jumping of genes that are used for curing genetic disorders. Jumping library allows the construction of met oncogene clones. As mammals cell consists of a number of repetitive DNA varieties that causes inappropriate cloning along with logjam in way of DNA replication. In respective aspects chromosome jumping has been found very appropriate as it allows following advantages:

- Permits speedy movement through the genome as compared to various other methods, like chromosome walking.
- Chromosome jumping allows formation of jumping library of un-cloneable regions into bacterial hosts.

- Technique also allows formation of genetic markers that have recognized chromosomal locations.
- Amalgamation of jumping with linking jumping libraries to walking allows probability of directional walking and might allow the analysis of longer regions in parallel mapping strategies.
- Decreases the complication for the formation of library screening as well as construction of mammalian genome.

Chromosome jumping allows adverse advantages; but yet there are still some restrictions in the cloning of vectors. More specifically for cloning of vectors which have hundreds of kilobase distance. Moreover, jumping also doesn't allow the cloning of prevailing DNA. Yet chromosome jumping is highly as it allows jumping of more than 100 kilobases (Leslie, 2008).

KEYWORDS

- **chromosome jumping**
- **chromosome library**
- **DNA analysis**
- **electrophoresis**
- **prenatal diagnosis**
- **prokaryotes and eukaryotes**

REFERENCES

- Collins, F. S., Drumm, M. L., Cole, J. L., Lockwood, W. K., Vande, W. G. F., & Iannuzzi, M. C., (1987). Construction of a general human chromosome jumping library, with application to cystic fibrosis. *Science*, 235(4792), 1046–1049.
- Drumm, M. L., (2001). Construction of chromosome jumping and linking libraries in *E. coli*. *Current Protocols in Human Genetics*, 1(1), 5.4.1–5.4.17.
- Leslie, A. P., (2008). Transposons: The jumping genes. Transposons: The jumping genes. Nature Education, 1(1), 204.
- Madireddy, A., & Gerhardt, J., (2017). Replication through repetitive DNA elements and their role in human diseases. *Advances in Experimental Medicine and Biology*, 1042, 549–581.

- Poustka, A., & Lehrach, H., (1986). Jumping libraries and linking libraries: The next generation of molecular tools in mammalian genetics. *Trends in Genetics*, 2, 174–179.
- Poustka, A., Pohl, T. M., Barlow, D. P., Frischauf, A. M., & Lehrach, H., (1987). Construction and use of human chromosome jumping libraries from NotI-digested DNA. *Nature*, 325(6102), 353–355.
- Pray, L., & Zhaurova, K., (2008). Barbara McClintock and the discovery of jumping genes (transposons). *Nature Education*, 1(1), 169.
- Sandeep, R., (2012). Barbara McClintock and the discovery of jumping genes. *Proceeding of the National Academy of Science in USA*, 109(50), 20198, 20199.
- Slade, I., Stephens, P., Douglas, J., Barker, K., Stebbings, L., Abbaszadeh, F., Pritchard-Jones, K., et al., (2009). Constitutional translocation breakpoint mapping by genome-wide paired-end sequencing identifies HACE1 as a putative Wilms tumor susceptibility gene. *Journal of Medical Genetics*, 47(5), 342–347.
- Slotkin, R. K., & Martienssen, R., (2007). Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics*, 8, 272–285.
- Talkowski, M. E., Ernst, C., Heilbut, A., Chiang, C., Hanscom, C., Lindgren, A., Kirby, A., et al., (2011). Next generation sequencing strategies enabling routine detection of balance chromosome rearrangement for clinical diagnosis and genetic research. *American Journal of Human Genetics*, 88(4), 469–481.

CHAPTER 10

Electrophoresis

MUHAMMAD ISHTIAQ, MAHNOOR MUZAMMIL, and
MEHWISH MAQBOOL

*Department of Botany, Mirpur University of Science and Technology
(MUST), Mirpur, Azad Jammu and Kashmir (AJK), Pakistan*

ABSTRACT

The present chapter covers the emergence and importance of electrophoresis as a tentative technique. Electrophoresis technique has found extensive adaptation by recent era owing to excellent analytic and bimolecular separation on the bases of their respective masses and charge density. Different types of electrophoresis are used for nucleic acid and protein analysis. Along with this electrophoresis has been found very efficient in the analysis and separation of metallic oxides and nanoparticles. This tentative technique works on electrokinetic principle and forms different light or dark, thin, or thick bands on the bases of charge to size ratio. The chapter covers the basis working principle, gel types, buffers, and different types of electrophoresis technique recently flourished for analysis and separation of different classes of biomolecules.

10.1 INTRODUCTION

Electrophoresis was first used by Greeks (ηλεκτροφόρηση) which means ‘electron bearing.’ Electrophoresis is the motion of discrete particles in

fluid containing different charges under uniform electric field (Lyklema, 1995; Dukhin & Goetz, 2017). The technique of electrophoresis is a potential way to analyze the macromolecules on the bases of their charge to size ratio. This physical method involves separation of those particles which are capable of acquiring charges. Positive charges are called cataphoresis whereas negative charged particles are called anaphoresis. Electrophoresis is actively used for the analysis, separation and study of DNA, RNA, and protein (Dukhin & Goetz, 2017).

10.2 HISTORY

Electrophoresis techniques were first proposed by Tiselius in 1937. The technique was put forward on the basic principle of Faraday's law of electrolysis in the late 18th century. Techniques evolved after progressive experimental work of Johann Wilhelm Hittorf, Walther Nernst, and Friedrich Kohlrausch put forth for the separation of smaller molecular ion through the aqueous solution. Electric field was provided to system as source of energy. Electric supply facilitated the movement of positive and negative charges towards cathode and anode hence demonstrating electrochemistry of aqueous solutions. Afterwards Kohlrausch presented the mathematical equation for the fluctuating concentrations of charged particles within the solution (Figure 10.1). Kohlrausch equation explained the remarked movement of particles with high-pitched across the boundaries on bases of charge difference.

After the successful work of Arne Tiselius on electrophoresis, the technique was adopted for biomolecules and chemical analysis as well as separation of molecules. Tiselius work lead under supportive hands of "Rockefeller foundation," facilitated later research on nucleic acid study, cloning of DNA, formation of markers and many more. Tiselius presented and apparatus in 1937 as Tiselius Apparatus in "A new apparatus for electrophoresis analysis of colloidal mixtures."

The electrophoresis technique was adopted gradually still the evolution of effective zones electrophoresis in 1940s and 1950s, in which filter paper and gel was used as supportive medium. Afterwards till 1960s, sophisticated gel electrophoresis method was adopted which enable researchers to distinctly isolate biomolecules on the bases of size to charge variances, that facilitated huge progressive turn over in molecular biology. Since

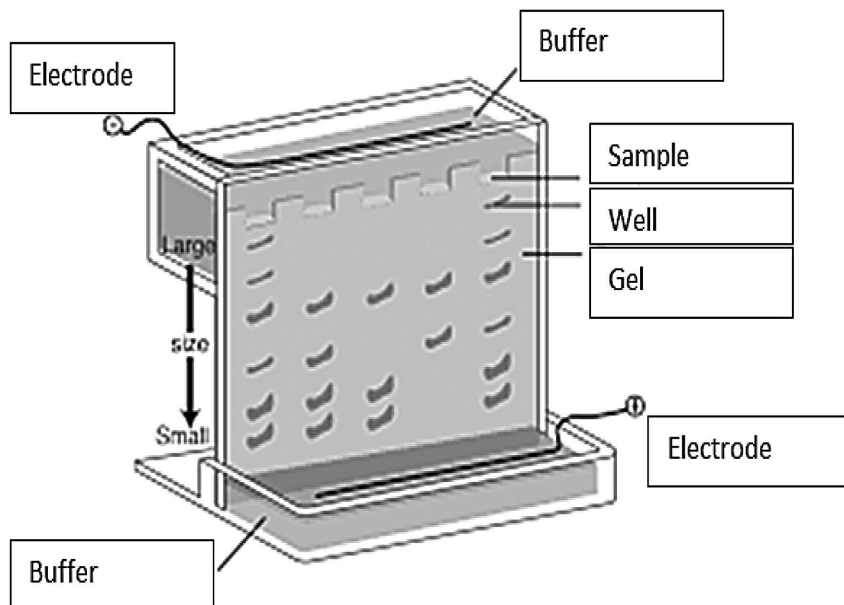


FIGURE 10.1 Electrophoretic mobility: Speed of moving particles per electric current intensity.

electrophoresis has been an important tool for the separation, analysis, and cloning of nucleic acid molecules. Electrophoresis open gateway for molecular analysis of a number of different biological techniques like protein fingerprinting, southern bolt and other blotting processes, DNA sequencing, and many more.

10.3 WORKING PRINCIPLE OF ELECTROPHORESIS TECHNIQUE

Electrophoresis is an essentially important technique now-a-day for the analysis and separation of biomolecules on the bases of charge to size ratio. Technique works on the Electro-kinetics' principle. Every biomolecule has a charge on depending upon its molecular composition. In electro-phoresis electric current forms an electrolytic cell in gel medium in which particles moves and separates on the bases of their charge to size ration in-between electrodes (Roth, 2005). Electro-kinetics uses electric current and facilitates the movement of particles on the bases of their charge and mass energies (Dukhin & Goetz, 2017).

Nucleic acid is negatively charged due to presence of phosphate groups. In nucleic acid different nitrogenous bases and phosphate linkage (double and single standard molecules) are involved due to mass to charge ratio of each nucleotide is different. Negative charge on phosphate gives acidic pH to the nucleic acid. Due to this charge variance on nucleic acid, it moves towards the cathode and anode on the onset of electric field. Electrophoresis technique requires some biased parameters to perform separating and analyzing nucleic acid. This is consummate via using sieving media like a gel that will facilitate the movement of molecules on the bases of their size (larger and smaller) and charge (higher or lower). As gel is viscous material so smaller particles with lower charges will separate earlier while larger molecules will face more fractions while passing through gel material. Smaller and sphere particles stabilize the electrical and viscous force hence creasing electrophoretic mobility* (u). U is equal to:

$$U = \frac{q}{6\pi R\eta}$$

In following equation q represents charge that a particle bears, while R is the radial distance of particle and η is viscosity of solvent. The movement of nucleic acid in gel electrophoresis also depends upon the shape of the molecules under test (Erlandsson & Robinson, 2011). High-order structure of nucleic acid will cause significant divergent from simple linear structure molecules. On some aspects electrophoresis is also used for the determination of the structure composition of DNA molecule. The external force exerts a potential on the diffusible that in result causes tangling motion in fluid in accordance to the charge on the contiguous surface. Impacting force can be via electronic, due to unequal pressure ascent, deliberation gradient or gravity. However, the movable phase can be constant fluid or diffusible phase (Lyklema, 1995; Dukhin & Goetz, 2017).

10.4 SUPERCOILED AND LINEAR DNA STRUCTURE

Double standard ds DNA, i.e., supercoiled molecule is much more similar to spiral DNA, while linear molecule parades polymeric nature by its regular movement. Linear DNA molecule is believed to appear in wriggling form, forming snake like pattern on electrophoresis gel medium. Conversely, the ds DNA causes lesser flexibilities in molecules structure resulting in the repetition of nucleic acid which formulates more rigidity at

joints and decreases the extent of movement of particles through the sieve gel. As a result of such a behavior, the ‘tail’ of higher molecular weighted DNA chain no longer samples is spread in the gel, and movement of DNA becomes independent of its length (Katarina et al., 2014). For the separation of DNA with greater (460 kbp) molecular weight other techniques like pulsed-field gradient electrophoresis is needed. Under a constant electric field, the movement of the particles is directly proportional to the well size in the gel medium.

10.5 GENERAL WORKING

The electrophoresis technique is an evolutionary invention in biological science that enables the separation of molecules (DNA, RNA, protein, and various organelles) on the bases of their size and charges. Electrophoresis consists of a tank supplied with electric field and gel in which test molecules are set to move according to charge to size ratio (Bengtsson et al., 2014). Generally, the technique uses either agarose or Polyacrylamide gel in electrophoresis chamber. Electric field supplies negative charge at one end that thrusts the molecules within the gel while positive charge charms molecules through the gel. Molecules are dispersed in gel via dispensing well in the gel material (Figure 10.2).

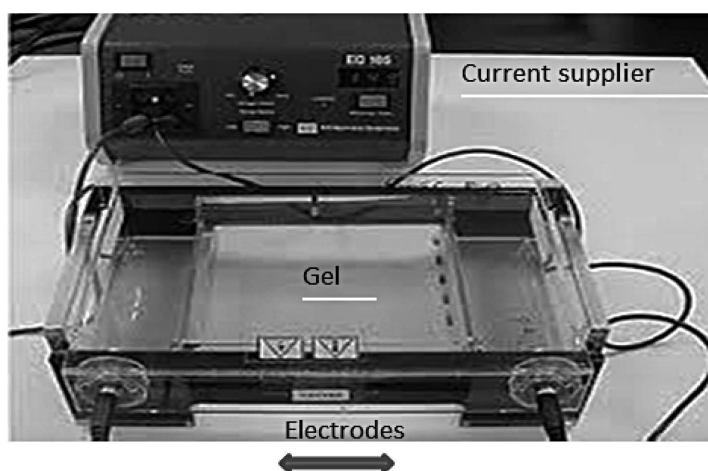


FIGURE 10.2 Gel electrophoresis apparatus used for study of protein electrophoresis process.

After loading the molecules in the well of the electrophoresis chamber, it is connected to the power source. Once on the onset of the electric supply, the molecules move according to the molecular masses and charges. Denser molecules move slowly while lighter molecules move faster (Boyer, 2000; Hofmann & Samuel, 2018).

The constancy and type of gel is selected with reference to the targeted material under test. The gel is a cross linked polymer that provides Electrophoretic mobility to the macromolecules under test. The electrophoretic movement of the molecules highly depends upon the length, conformation, and charge of the test molecule. While separation of protein or small Nucleic acid (DNA, RNA, and oligonucleotide) gel is generally composed of varying concentration of acrylamide and a cross-linker. They form lattice network of varying size of polyacrylamide. For the separation of higher molecular weight (more than 100) nucleic acids purified agarose gel is usually used. Agarose belongs to extensive un-branched chained carbohydrates with neutral charge. Agarose gel lacks cross-links which allow larger pores formation in the gel chamber, enhancing separation of macromolecules and the efficiency of the electrophoresis process. In comparison Acrylamide, is a 'Neurotoxin' which requires appropriate precautionary measures for evading poisoning of gel while polyacrylamide gel don't require such precautionary measures (Almdal et al., 1993; Khademhosseini & Demirci, 2016). Polyacrylamide gel electrophoresis (PAGE) is more specifically utilized for separating and analyzation the RNA molecules. Both gels form solid and porous matrix in the electrophoresis process (Petrov et al., 2013).

10.6 ELECTROMOTIVE FORCE (EMF)

During the process of electrophoresis electric current provides electromotive force (EMF) to the gel matrix. Macromolecules are placed within the well made in gel. EMF causes movement of the macromolecules across well in gel according to their charge to size ratio. The rate of movement of different molecules is different that is measured from their respective masses. Charge to size ratio (Z) of all test species are same for the uniform movement of molecules, whereas when the e/m value of all the test species is different molecules move according to their respective e/m value. When electric current is provided to the gel chamber, it establishes an electrolytic

cell in which net positive charge moves towards the anode while negative charge moves towards the anode, i.e., positive charge. Gel chamber provides the area for the separation and movement of non-uniformly charged test species toward their respective electrodes (Khan, 2013).

When different samples are loaded in the electrophoresis gel medium, then each particle will run in a parallel pattern in separate lane. In mixture, different particles show distinct bands depending upon the charge and size of separating particles. Complete separation of particles shows different bands distant from the original mixture while incomplete separation of particles shows overlapping of bands. While particles having the same size and charge are represented as multiple uncertain components. The separated bands that fall at the same line and distance from the top usually have the same size and charge. Electrophoresis is sometimes used for the analysis of different particles. Electrophoresis helps in recognizing the size and charge of unknown particles for this purpose, molecular markers of known molecular weight are used. Markers of known molecular weight are run with parallel to unknown particles; representing bands are compared with known values. The size to charge ratio of molecules is compared with the distance traveled by respective molecules (Swathi et al., 2017).

10.7 BUFFERS

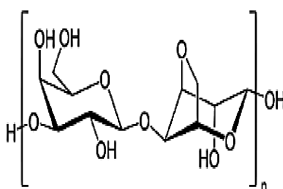
Gel being a sensitive material has specific electrophoretic efficiency. When constant current passes through the gel matrix it causes heating of the gel, thus melting gel during electrophoresis. To avoid melting of gel and constant process of electrophoresis certain buffers are added in gel matrix solution to maintain the pH of the gel. The charges on macromolecules such as DNA and RNA are pH sensitive. Long term electric supply could alter the pH of solution so buffers maintain the charges on DNA and RNA (Minde, 2012). Along with this, during the determination of molecular mass of unidentified protein via sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), buffers are extremely important to restrain the original charge on unknown protein. Various buffers are used in electrophoresis depending on type of macromolecule. For example: Tris/acetate/EDTA (TAE), Tris/Borate/EDTA (TBE) is used for nucleic acid analysis (Barasinski & Garnweitner, 2020). Since different buffers are available for example: lithium borate LB buffer, isoelectric histidine,

etc. Sometimes low current is used in electrophoresis gel medium that increases the shelf life of buffers solution. Such buffers don't affect the mobility of the particles through the gel medium. While dealing with RNA molecules analysis and separation borate-based buffers are not used as they polymerize and intermingle with cis diols commonly present in RNA molecule (Ninfa et al., 2009). TAE has low buffer capacity but is generally used for DNA analysis and separation as it provides better resolution power for the large DNA molecule. This enhances the time duration of the process due to lower rate of current supply without affecting the results. After the evolution of LB buffers TAE was replaced as LB buffers provides better results at high voltage and reaction rate (35 V/cm). Along with this LB buffers have higher resolved power and can analyze DNA fragments with up to 5 kbp length (Brody & Kern, 2004).

10.8 DIFFERENT GEL TYPES IN ELECTROPHORESIS

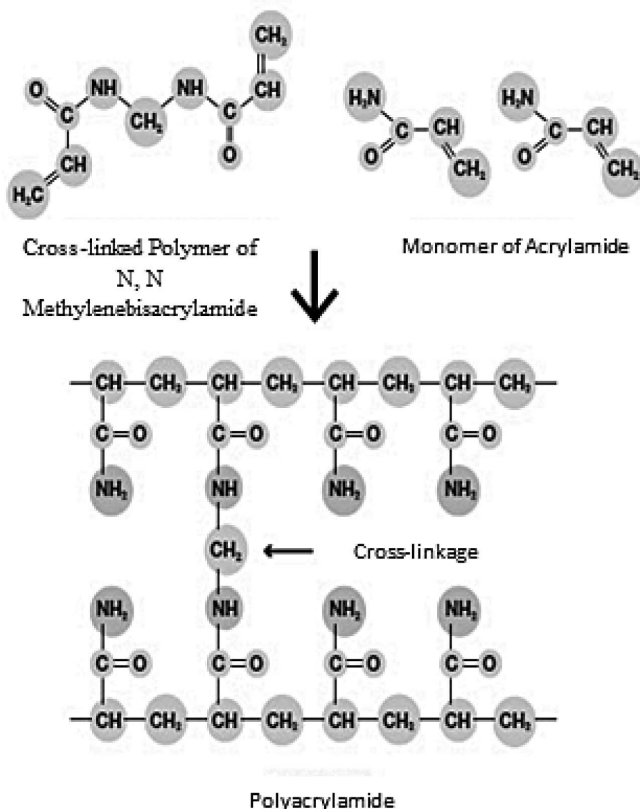
Generally, two (agarose and polyacrylamide) types of gel medium are used in electrophoresis process depending upon the nature of particle under test. Gel is selected in accordance to size and charge density of analyst.

10.8.1 AGAROSE GEL



Agarose gel is a natural extract from either seaweed or from red algae. Agarose consist of extensive linear chains of d-galactose along with 3,6-anhydro- α -L-galactopyranose. Agarose gels are purified at higher extents before being used in the electrophoresis process. This purification adds extra florescent abilities and removes all types of impurities from gel that would react with DNA. Agarose gel is generally preferred for analyzing and separation of DNA molecules up to 50–20,000 base pairs in length. There are several advantages of agarose for DNA molecule analyzes such as:

- Purified agarose gel medium increases visualization of DNA molecule bands;
- Agarose gel also shows excellent amplification power for DNA molecule of 50 bp to megabases (millions of bases), that generally requires separate setup;
- Agarose gel set more easily as compare to other gel medium and show zero to minimum chemical interaction with test particles as it set in semi-solid physical state;
- One completion of the process gel could be frozen and reused. After the experiment is finished, the resulting gel can be stored in a plastic bag in a refrigerator;
- Agarose gels are also the best medium for the analysis of large protein molecules with sizes up to 200 kDa.



The length difference between two adjacent bands of DNA fragments is directly proportional to the percent viscosity of agarose gel. A higher percentage of agarose gel will require more flow time for molecules across the gel. More concentrated agarose gel requires pulsed-field gel electrophoresis (PFE), or Field inversion electrophoresis. For the resolving DNA with 5–10 kilobases length generally agarose gels is made with 0.7% concentration while concentration is increased to 2% for resolving DNA with smaller length, i.e., 0.2–1 kilobase along with buffers. Whereas, 3% concentration is used for extreme smaller DNA fragment using polyacrylamide gel. Gel with lower concentrations causes breakage when subjected to higher degrees of lifting. On the other hand, gels with higher concentrations are habitually fragile and shows uneven distribution. So generally, gel with around 1% concentration is used for most purposes.

10.8.2 POLYACRYLAMIDE GEL ELECTROPHORESIS (PAGE)

Polyacrylamide gel electrophoresis (PAGE) is a general technique implied for the analysis and separation of the proteins and amino acids with an average size of 5 to 2,000 kDa. Polyacrylamide gel was more specifically used for protein analysis as gel forms well of uniform size. Pore formation in gel medium is managed by adjusting the concentration of the acrylamide and bis-acrylamide implied in the formation of the gel medium. Due to neurotoxic nature (in liquid form) of acrylamide, it requires greater care while its preparation. Generally, DNA sequencing techniques such as “Maxam–Gilbert or Sanger sequencing” are used which utilizes polyacrylamide gels for the separation of DNA fragments. Generally, for the separating DNA fragments, agarose gel is used, but for smaller DNA fragments, polyacrylamide gel is used. Along with separation of smaller DNA fragment, polyacrylamide gel is also applicable in the field of immunology along with protein analyzation. In addition, polyacrylamide gel is used for separating various protein molecules and isomers of proteins in distant bands.

Polyacrylamide gel concentration used in electrophoresis highly depends upon the length of protein and probes length, i.e., higher the length of protein higher the concentration of gel. Classically 6%, 8%, 10%, 12% or 15% of polyacrylamide gels is used in the electrophoresis process. Changing the concentration of buffer solution adds more resolution power to the protein analyzation (Schägger, 2006).

10.8.3 GEL CONDITIONS

10.8.3.1 DENATURATION

Sometimes denaturizing gels are used for destructing the structure of test particles. Denaturing gel converts the spiral structure into a linear structure. This is required for increasing the mobility of biomolecules which depends on straight structure and charge to mass density. Consequently, denaturizing gel causes the destruction of the secondary, tertiary, and quaternary structure of molecules parting only the primary configuration of molecules for analyzation.

For the denaturing of the biomolecules (DNA and RNA) under test in electrophoresis is useful in band formation accomplished by temperature gradient gel electrophoresis (TGGE) and denaturing gradient gel electrophoresis (DGGE) (Fischer & Lerman, 1979; Zilberstein et al., 2007). RNA molecules are compulsorily denaturized as it forms more intermolecular interaction as compared to DNA that reduces the mobility of Electrophoretic. In respective cases urea, DMSO and glyoxal are used as denaturizing agents for destructing the ultra-RNA structure. Some other means of denaturation methods implied in electrophoresis are:

- For the denaturizing of the nucleic acids generally urea is added along with buffer solution;
- For protein denaturing sodium dodecyl sulfate (SDS) is used;
- For complete denaturalization of proteins, the disulfide covalent bond present between different amino acids is destructed by using reducing PAGE method (Wittig & Schägger, 2005);
- Reduction is accomplished by beta-mercaptoethanol or dithiothreitol;
- Broad-spectrum analyzation of protein is accomplished via reducing PAGE;
- Traditionally, high lethal methylmercury hydroxide is used as denaturing agent for RNA in the electrophoresis process.

10.8.3.2 NATIVE

For the maintenance of the ultra-structure of biomolecules (testing molecule) native gels are used that prohibits the denaturizing of biomolecules

under test. Native gel is generally used for the restoration of the natural bio-structure of molecules either: primary, secondary, tertiary, or quaternary. This original structure restoration causes lower mobility of the electrophoresis process. Typically, native gel electrophoresis is utilized for Proteomics and Metallome. Alongside native PAGE is also applied for scanning the genes (DNA) that is helpful for the detection of mutations such as present in Single-strand conformation polymorphism. Native gel is also used for the identification and presence of enzymes while purification of proteins.

Likewise denaturizing, native gel electrophoresis doesn't use any denaturizing agent. Molecules are generally separated as either protein or as Nucleic acids. Due to which they not only differ in chemical properties but also differentiates at physical levels (charge and mass). Hence protein bands are visualized by adding protein staining along with enzyme staining.

10.8.3.3 VISUALIZATION

Staining is an important part of the electrophoresis process as staining makes band of different particles more visible. Hence biological staining is of prime importance in the electrophoresis process. For visualization of different biomolecules following staining agents are used:

- For the visualization of DNA more typically ethidium bromide is used. Which, when intercalated into DNA, fluorescence under ultraviolet light, while;
- For protein visualization silver staining or Coomassie brilliant blue dye is used;
- For separation of radioactive molecules in DNA sequencing gel, isotopic tracers are taken from gel or sometimes photograph of gel bands are taken using the Gel Doc system.

10.8.3.4 DOWNSTREAM PROCESSING

After the separation of different bands on gel they are downscaled for analysis the separated portions. Downscaling is a superfluous separation technique applied either by isoelectric focusing (IEF) or SDS-PAGE. The

multiplexes proteins extracted on gel are cutoff physically. Each portion is analyzed by Peptide mass fingerprinting or De novo peptide sequencing after In-gel digestion. These techniques provide vast information about the ultra-structures of extracted biomolecules.

10.9 TYPES OF ELECTROPHORESIS

Innovated electrophoresis techniques were revolutionized by the late 1940s. New innovation in the electrophoresis process overcomes the short comes and extended the boundaries of Tiselius apparatus. Previously broad-spectrum separation of biomolecules was not possible through electrophoresis but later progress in technique made many of flourishment. Older electrophorese used movement of freely moving charged particles via solution, while later introduce methods utilized gel or other solid medium for the separation of particles in distinct and firm bands (zones). Broadly electrophoresis is divided into two groups:

1. **Zone Electrophoresis:**
 - i. Paper electrophoresis;
 - ii. Gel electrophoresis;
 - iii. Thin layer electrophoresis;
 - iv. Cellulose acetate electrophoresis.
2. **Moving Boundary Electrophoresis:**
 - i. Capillary electrophoresis (CE);
 - ii. Isotachophoresis;
 - iii. Isoelectric focusing;
 - iv. Immune electrophoresis.

10.9.1 ZONE ELECTROPHORESIS

Zone electrophoresis is most commonly used for the analysis and separation of the proteins, nucleic acids, and polymers. Technique was introduced by Oliver Smithies around 1955s. In this technique test molecules are dissolved in a specific buffer solution and ran under constant electric supply. For example, for protein separation acetic acid and tris-hydroxymethyl aminomethane is commonly used. Different particles set in wells

according to their mobility which is directly proportional to the active masses and charges on particles. In resultant graph peaks of different length and variability describes the percentage of varying amino acids constituting the proteins. Zone electrophoresis generally uses starch gel “Polyacrylamide” along with other gel combination that allows separation of proteins with minor differences in structure, masses, and charge density. For clinical trials and analysis of isoenzymes 0.7–1% agarose gel is used that provides efficient analysis of protein serum and lactate dehydrogenase (Ldh) along with creatine kinase, respectively.

10.9.2 GEL ELECTROPHORESIS

Gel electrophoresis is generally used for the separation of proteins and nucleic acid (DNA and RNA). These macromolecules are separated on the bases of charge density and masses. Gel electrophoresis is also actively used for the separation of a mixture of unknown protein and nucleic acid on bases of mass to charge ratio.

Gel electrophoresis used constant electric supply that forces the movement of biomolecules through gel medium (either agarose or any other). As nucleic acid is negatively charged, the electric field causes movement of positive and negative charge particles towards the anode and cathode. Molecules with smaller size moves more efficiently through gel medium while those with higher molecular weight moves slowly. Respective movement of molecules is referred as sieving. Agarose gel is used for the separation of proteins as agarose gel separates particles on bases of charges where pores in the medium are too small for protein sieving. Gel electrophoresis is best suited for nanoparticles parting.

Gel electrophoresis is widely used owing to anti-convective nature of gel medium or sieving medium. As gel suppresses the thermal conventional force created due to electric supply and acts as a sieving medium, hindering the path of biomolecules. Along with this gel also provided an ideal medium for staining of extracted particles (Ott et al., 2010).

10.9.3 THIN LAYER ELECTROPHORESIS

Thin layer electrophoresis lost its importance after the development of more feasible gels like agarose and polyacrylamide gels. These gels have a

high capacity of loading molecules and show better and enhanced separation and analysis power. TLE is mainly confined with analysis of high molecular weight polysaccharides and lipopolysaccharides. Respected molecules are analyzed via TLE as pores formed in silica gel hinder high molecular weight particles. Process uses thinner layered silica gel medium along with buffers of high molecular, which would obstruct the pores of the gels (Scherz, 1990).

10.9.4 CELLULOSE ACETATE ELECTROPHORESIS

Cellulose acetate is obtained by treating acetate salt of cellulose extracted from cotton treated with acetic acid in presences of catalyst, i.e., Sulfuric acid. Particles migrate on the buffer film at the surface of the cellulose acetate plate or membrane. Separation of particles (amino acid) is accomplished on the bases of charge density. Cellulose is used in electrophoresis for the analysis of different variants, along with elution of quantity of hemoglobin's, A2, A, S, D, Lepore, α -chain variants, Hb H and Hb Bart's.

Cellulose acetate electrophoresis consists of 2–3 acetyl groups per glucose unit with lesser adsorptive abilities as compared to paper electrophoresis yet providing sharp bands against bright background for staining. Celluloses acetic acid technique is commonly used for clinical and biological protein sample (albumin and globulins).

10.9.5 CAPILLARY ELECTROPHORESIS (CE)

Capillary electrophoresis (CE) has found its importance in recent era owing to its analytic abilities for micropreparative (Jorgenson & Lukacs, 1981; Hjertén, 1983), HPLC (high-performance liquid chromatography), and HPCE (high-performance capillary electrophoresis). Molecules are separated using fused silica capillaries with tube length of 20–30 cm and diameter around 50–100 μm . Process involves submergence of ends of capillary tubes in buffer. The medium is supplied with electrodes. Process requires very chemicals and sample in very smaller amount, i.e., only 2–4 nl nanogram of testing material is required for proceeding the process. During the process current supply is maintained to 1 kV cm^{-1} and currents of 10–20 mA are generally utilized. Due to this only 30 kv of current is required as an energy source. Using a fan heat can be induced

into the system via the thin capillaries. In CE typically separation takes up to 10–20 mins. A number of different methods are available for the detection of particles some of them are: UV-Vis, fluorescence, conductivity, electrochemistry, etc. Generally, fraction analysis is achieved via UV measurements at around 280, 260 while in some cases up to 185 nm using direct capillaries are used. Certain identification analysis requires lower detection limits, i.e., at the attomole level. After this the resultant particles are analyzed HPLC interpreted using software. As a precautionary measure, polyacrylamide or methylcellulose are coated on the inner side of capillary tubes to avoid molecules adsorption. CE can be implied at all separation levels, i.e., electrophoresis, isotachopheresis (ITP), and IEF. Different molecular analysis requires varying buffers combinations for example: 20–30 mmol l⁻¹ sodium phosphate buffer, pH 2.6, for peptide electrophoresis.

10.9.6 ISOTACHOPHORESIS (ITP)

For the analysis and separation of ionic biomolecules isotachopheresis (ITP) is used. It is an analytic analysis of molecules chemistry utilized for separating some specific ions from a mixture of solution. ITP reflects the active movement of the ions through an electric field that depicts the migrating power of ions in a medium.

10.9.7 MICROCHIP ELECTROPHORESIS (MCE)

In planar process MCE is generally a reduced form of CE that allows higher mechanization and faster analyzation. Using narrow channels, the particles are separated and analyzed. These particles are then introduced on glass, silicone, or polymer surfaces by photolithographic process. Generally, the size ranges from 1–10 cm with about 50 µm diameters. Medium tank is connected with electric supply and electrodes are established that provide room for the injection and separation. About 1/10th CE are necessary for the ~0.1–0.5 nl, injection of electro kinetic energy, with electric voltage supplied at 1–4 kV. The separation time duration of MCE is 4 times lesser than that of CE, which requires 50–200 s. The most commonly used detection method is laser-induced fluorescence (LIF) as, for instance, described by Schulze et al. (2010).

10.10 CONNOTATION

- Electrophoresis is widely used for the size estimation of DNA molecules, e.g., in Restriction enzyme of cloned DNA.
- It has also found its importance in analyzation of PCR products, e.g., in molecular Preimplantation genetic diagnosis or Genetic fingerprinting.
- Electrophoresis also enables us to separate restricted genomic DNA prior to Southern blot of RNA prior to Northern Blot.
- Gel electrophoresis is utilized in fields of Forensic, fingerprinting, DNA analysis, antibiotic formation, markers formation, etc. (Strecker et al., 2010).
- Resultant bands are visualized via by ultraviolet light or gel imaging device. Photographic version is detailed via computer-operated camera. Different bands thickness and intensity is noticed and compared with the standard markers using specific software (Schulze et al., 2010).
- Through electrophoresis defects of Nucleic acid are pinpointed. Different mutations in RNA that could lead to serious diseases formation or can lead to any genomic abnormality are detected by electrophoresis. For example, normal eukaryotic genome shows distinct RNA bands of 28s and 18s rRNA, while mutant RNA shows less sharp bands with 2:1 intensity (Lodish & Matsudaira, 2004).
- Electrophoresis is also actively involved in proteins analysis and separation.
- For proteins analysis special coating of detergent like: Sodium dodecyl sulfate (SDS) is used that provides negative charge to protein molecules as generally proteins are neutrally charged. Negative charge spread on protein causes active movement of charges towards cathodes and anode and the extent of their speed reflects their charge to size ratio (Lodish et al., 2004).
- Along with this electrophoresis also provided denaturation of 3D structure of protein without effects the biochemical nature of protein molecules. Denaturizing agents only degenerates the disulfide bonds between spiral or collided chain and convert them in long linear chain for steady mobility across gel medium.

- So, protein is generally analyzed using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), by Native Gel Electrophoresis, by preparative gel electrophoresis QPNC-PAGE, or by electrophoresis characterization via ligand interface generally accomplished by electroblotting or by affinity electrophoresis in agarose or by CE as for estimation of Binding constant and determination of structural features like Glycan content through Lectin binding (Gordon, 1975; Perrett, 2007).
- Another novel solicitation of electrophoresis is the separation of metallic oxides and nanoparticles (e.g., Au, Ag, ZnO, and SiO₂). Particles are separated on the bases of the charge density, masses, and biochemistry of nanoparticles. The scope is to obtain a more homogeneous sample (e.g., narrower particle size distribution), which then can be used in further products/processes (e.g., self-assembly processes) (Barasinski & Garnweitner, 2020).

KEYWORDS

- **bands**
- **biomolecules**
- **capillary electrophoresis**
- **electrokinetics**
- **electrophoresis**
- **isotachophoresis**
- **nucleic acid**
- **proteins**

REFERENCES

- Almdal, K., Dyre, J., Hvidt, S., & Kramer, O., (1993). Towards a phenomenological definition of the term ‘gel’. *Polymer Gels and Networks*, 1(1), 5–17.
- Barasinski, M., & Garnweitner, G., (2020). Restricted and unrestricted migration mechanisms of silica nanoparticles in agarose gels and their utilization for the separation of binary mixtures. *J. Phys. Chem. C*, 124, 5157–5166.

- Bengtsson, K., Nilsson, S., & Robinson, N. D., (2014). Conducting polymer electrodes for gel electrophoresis. *PLoS ONE*, 9(2), e89416.
- Boyer, R. F., (2000). *Modern Experimental Biochemistry* (3rd edn.). US: Pearson.
- Brody, J. R., & Kern, S. E., (2004). History and principles of conductive media for standard DNA electrophoresis. *Anal. Biochem.*, 333(1), 1–13.
- Dukhin, A. S., & Goetz, P. J., (2017). *Characterization of Liquids, Nano- and Micro-Particulates and Porous Bodies Using Ultrasound*. Elsevier.
- Erlandsson, P. G., & Robinson, N. D., (2011). Electrolysis-reducing electrodes for electrokinetic devices. *Electrophoresis*, 32, 784–790.
- Fischer, S. G., & Lerman, L. S., (1979). Length-independent separation of DNA restriction fragments in two-dimensional gel electrophoresis. *Cell*, 16(1), 191–200.
- Gordon, A. H., (1975). *Electrophoresis of Proteins in Polyacrylamide and Starch Gels*. New York: American Elsevier Publishing Company, Inc.
- Hofmann, A., & Clokie, S., (2018). *Wilson and Walker's Principles and Techniques of Biochemistry and Molecular Biology*. Great Britain: Cambridge University Press.
- Khademhosseini, A., & Demirci, U., (2016). *Gels Handbook: Fundamentals, Properties and Applications*. World Scientific Pub Co Inc.
- Khan, M. F., (2013). Online measurement of mass density and viscosity of pL fluid samples with suspended microchannel resonator. *Sens. Actuators, B*, 185, 456–461.
- Lodish, H., Berk, A., & Matsudaira, P., (2004). *Molecular Cell Biology* (5th edn.). WH Freeman: New York, NY.
- Lyklema, J., (1995). *Fundamentals of Interface and Colloid Science*, 2, 3, 208.
- Minde, D. P., (2012). Determining biophysical protein stability in lysates by a fast proteolysis assay. *PLOS One*, 7(10), e46147.
- Ninfa, A. J., Ballou, D. P., & Benore, M., (2009). *Fundamental Laboratory Approaches for Biochemistry and Biotechnology* (p. 161). Hoboken, NJ: Wiley.
- Perrett, D., (2007). From 'protein' to the beginnings of clinical proteomics. *Proteomics: Clinical Applications*, 1(8), 720–738.
- Petrov, A., Tsa, A., & Puglisi, J. D., (2013). Chapter 16 – Analysis of RNA by analytical polyacrylamide gel electrophoresis. In: Lorsch, J., (ed.), *Methods in Enzymology* (Vol. 530, pp. 301–313). Academic Press.
- Roth, C. M., (2005). Electrophoresis nucleic acids. *Encyclopedia of Analytical Science*, 456–460.
- Schägger, H., (2006). Tricine-SDS-page. *Nature Protocols*, 1(1), 16–22.
- Schulze, P., Link, M., Schulze, M., Thuermann, S., Wolffbeis, O. S., & Belder, D., (2010). A new weakly basic amino-reactive fluorescent label for use in isoelectric focusing and chip electrophoresis. *Electrophoresis*, 31, 2749–2753.
- Swathi-Chaudhari, Kamalesh-Chaudhari, Seokbeom-Kim, Faheem-Khan, Jungchul-Lee, & Thomas-Thundat. (2017). Electrophoresis assisted time-of-flow mass spectrometry using hollow nanomechanical resonators. *Scientific Report*, 7, 3535.
- Tiselius, A., (1937). A new apparatus for electrophoretic analysis of colloidal mixtures. *Transactions of the Faraday Society*, 33, 524–531.
- Wittig, I., & Schägger, H., (2005). Advantages and limitations of clear native polyacrylamide gel electrophoresis. *Proteomics*, 5, 4338–4346.
- Wraxall, B. G. D., & Culliford, B. J., (1968). A thin-layer starch gel method for enzyme typing of bloodstains. *J. Forensic Sci. Soc.*, 8(2), 81, 82.

Zilberstein, G., Korol, L., Antonioli, P., Righetti, P. G., & Bukshpan, S., (2007). SDS-PAGE under focusing conditions: An electrokinetic transport phenomenon based on charge neutralization. *Anal. Chem.*, 79, 821–827.

CHAPTER 11

Genetically Engineered Microorganisms

KRINA MEHTA,¹ ROHIT PATEL,¹ SAMEERA SHARMA,¹
ARPIT SHUKLA,² DWEIPAYAN GOSWAMI,¹ MEENU SARAF,¹ and
PARITOSH PARMAR³

*¹Department of Microbiology and Biotechnology, Gujarat University,
Ahmedabad, Gujarat, India*

*²Department of Biological Sciences and Biotechnology,
Institute of Advanced Research, University of Innovation,
Koba Institutional Area, Gandhinagar, Gujarat, India*

*³Department of Biotechnology and Bioengineering, Institute of
Advanced Research, Koba Institutional Area, Gandhinagar,
Gujarat, India*

ABSTRACT

Genetic engineering develops at an expeditious rate owe to the successful researcher work in the varied field of application. Genetically engineered microorganisms (GEM) are constructed by genetically altering the genome, either incorporating the desired gene or deleting the undesirable gene through various engineering technologies. GEM depicts wide ranges of applications in food and fermentation industries, bioremediation, biotech, therapeutic agrochemical field. The rapid development in biotechnology in combination with genome editing technology leads to the better genetic understanding and the development of uniquely engineered microorganism. A number of

genome editing techniques, including CRISPR/Cas technologies, are significantly being applied for engineering desired microbial strains. This chapter discusses genetic and metabolic strategies involved in the construction of modified microbes and their significant implementation over various fields. Besides, the successful incorporation of genetically modified (GM) or engineered microorganisms for the effective mitigation of pollutants via efficacious remediation methods and approaches is well discussed.

11.1 INTRODUCTION

From a decade, microorganisms are widely used in various fields of applications. The movement of genetic material between entirely different species owes to the advancement in modern biotechnology. Genetically engineered microorganisms (GEM) is desirably created amended with modified genomic composition by the addition of the preferred genome artificially. Animals, plants, microorganisms, either prokaryotic or eukaryotic, are successfully genetically altered with the new desired traits. Genetic engineering is typically incorporated to aid the production of desired traits and minimize the undesirable trait or block the factor that interferes with an elevation of desirable components. Many series of steps, such as selection, designing, construction, incorporation into the host, are encountered in the genetic modification of microorganisms (Batt, 2014). An incredible advantage is offered by the advancement of genetic engineering in food microbiology, environmental applications, and the therapeutic sector. The feasible alternative for food globalization and the increased agriculture market is the genetic enhancement of the microbial strains employed in the food industries. An extra dimension of food safety and quality is assured by the food prepared from the genetically engineered (GE) microbes with no harm to human health and the environment. GEM offers personalized therapeutic drugs based on the genetic predisposition, novel drug formulation with targeted actions (Peter et al., 2019).

Strict government regulations and public concerns accompany this emerging technology. Risk assessment is a necessary decision-making process in risk management regarding the use GMO constructed by transgene technology and its exposure in the ecosystem (Peter et al.,

2011; Prakash et al., 2011). The implications of genome alteration may not be limited to the particular traits of the substituted gene; thus, extra precautions are mandatory and it is vital to ensure that it will not affect the ecosystem and human population (Prakash et al., 2011).

11.2 METHODS FOR GENETIC AND METABOLIC MODIFICATION OF MICROORGANISMS

Over the years microbes have been manipulated to produce metabolites and products of interest. Furthermore, microbes have emerged as a substitute for chemical synthesis processes in the form of microbial cell factories (MCF). As MCF microbes can be utilized to perform multistep catalysis as well as *de novo* synthesis of metabolites of interest (Fisher et al., 2014). Selection of microbes for metabolic engineering is ideally determined by the type of end product whether the pathway is native to the organism or not. In case if microbe possesses the pathway in question, it can be further optimized using metabolic and genetic engineering approaches otherwise *de novo* biosynthetic pathways can be installed. Such pathways might be of organisms that are difficult and impractical to cultivate such as fastidious microbes or higher eukaryotic organisms (Fisher et al., 2014).

Metabolic engineering approaches often come into play once pathways are installed in the target organism. One of the most frequently utilized approaches in metabolic engineering is overexpression of desired pathway still this approach often provides suboptimum results when implemented unaided. Such negative results might be produced due to adverse effect on cell growth as most of the cellular energy directed to the expression and formation of overexpressed proteins (Li et al., 2014). Metabolic engineering can be further utilized to detect the pathways that are not required for cell growth as well as target product formation and the metabolic flux of the system can be directed towards the target gene via creating knockout mutants of such unwanted genes. Such knock mutants sometime lack stability which again impacts the growth rate of the organism. For such cases, small regulatory RNAs (sRNA) can be utilized to decrease the availability of transcripts for translation. To identify the targets for metabolic flux management in microbial systems number of computational tools have been developed (Na et al., 2013). All this tool identifies and

computes the metabolic flux of systems based on different parameters. OptKnock was initially developed to determine the genes that can be knocked out to increase the final product formation (Burgard et al., 2003). Further improvisation resulted in formation of OptForce which is capable of calculations pertaining to overall gene regulation (Ranganathan et al., 2010). To further solidify the strategies for metabolic engineering tools such as OptORF and k-OptForce were developed that can respectively incorporate transcriptional regulatory networks and kinetic constants in their computations to produce metabolic strategies that are more accurate and effective (Chowdhury et al., 2014; Kim & Reed, 2010). Apart from this, flux balance analysis with flux ratios (FBrAtio) is also available that can calculate the impact of enzymes and cofactor availability along with reaction thermodynamics on end product production. This tool identifies the critical branching points in the pathway that can have major implications on product and suggest the various genetic strategies that can be utilized to increase the production (McAnulty et al., 2012; Yen et al., 2013).

To accomplish the modifications suggested by various metabolic engineering tools number of genetic engineering techniques can be utilized that have been discussed in this section. A number of strategies have been employed to create genetically modified organisms (GMO) and MCFs, but the basis of all those methods remains the same that are up and down-regulation of the target gene, gene sequence modification via either addition or removal of the specific gene (Han, 2004). Strategies such as horizontal gene transfer (HGT), mutagenesis, DNA shuffling, and CRISPR-Cas system has been utilized for GMO creation via utilizing afore mentioned phenomena.

11.3 GENETIC ENGINEERING STRATEGIES

Horizontal gene transfer (HGT) among microbes has been one of the most widely employed strategies for the genetic engineering. Naturally, transformation, transduction, and conjugation play cardinal role in HGT. Among these three strategies for HGT, transduction, and conjugation are having very limited application due to certain limitations (Johnsborg et al., 2007). Transduction mainly suffers due to the host specificity of virus that carries the gene inside the host which restricts their application to the specific host (Shirshikova et al., 2016). On the other hand, conjugation

requires specific genetic elements that are not always available within target microbes and the control over the whole process remains elusive (Davison, 1999). Whereas transformation remains one of the leading strategies as it is relatively easier to attain and a number of different methodologies have been developed to cause the transformation of cells that includes the chemical, protoplast, electroporation, biolistic, and shockwave mediated transformation. In chemical transformation, utilization of calcium ions along with heat-shock treatment has been known for a while now. In recent years, carbonyl cyanide m-chlorophenyl hydrazone (CCCP) has been utilized to increase the competence of cell without the aid of heat shock (Panja et al., 2006). Furthermore, PEG and dimethyl sulfoxide (DMSO) along with lower temperature has been successfully utilized to facilitate the chemical transformation of *Escherichia coli* (*E. coli*). Protoplast of various bacteria and fungus can be created for transformation using lysing enzymes (yatalase, vinoTaste Pro, novozyme, etc.), as well as physical pressure of grinding and shock where ideal method depends on type of cells (Son & Park, 2020). Once protoplast is formed utilization of chemical transformation agents such as PEG and CaCl_2 can provide effective transformation in bacterial culture such as *Streptomyces fradiae* and *S. griseofuscus* but the viability of fungal mycelium shows negative impacts as a protoplast (Hopwood, 1981; Stewart & Stewart, 2017). Better transformation and stable viability of fungal mycelium has been observed with transformation conducted using biolistic transformation method (Rivera et al., 2014). Moreover, electroporation shows better transformation efficiency than chemical transformation. Here, cellular construction of microbial cells such as their cell wall and cell membrane complexity, number, and variety of endonucleases (ER) and exonucleases in target species plays a very crucial role. Both fungus and bacterial cells have been transformed using this methodology namely they are: *Clostridium thermocellum*, *B. anthracis*, *B. cereus*, *E. coli*, *Hemophilus pleuropneumoniae*, *Klebsiella pneumoniae*; *A. oryzae*, *A. niger* (Davis et al., 2000). Recently, shock wave mediated transformation has been utilized for the transformation of a number of diverse bacteria such as *Salmonella typhimurium*, *Pseudomonas aeruginosa*, *Helicobacter pylori* and *Mycobacteria sp.* Even though this methodology produces effective transformation even on spores major drawback of expense holds it back (Rivera et al., 2014).

Apart from the addition of exogenous DNA, modification of native genomic DNA of organisms via mutagenesis has played a very vital role in the development of GMOs. Here, various methodologies such as physical mutagenesis, chemical mutagenesis, transposition mutagenesis, site-directed mutagenesis, etc., have been employed. These methodologies of mutagenesis can ultimately result in either base substitution, deletion or addition in genomic sequence leading to an inheritable change in the character of an organism. Induction of A:T $\rightarrow$ T:A transversion and G:C $\rightarrow$ T:A transversion has been observed within microbes when exposed to X-ray whereas UV radiation causes pyrimidine dimers (Shibai et al., 2017). Apart from UV and X-ray, gamma, and boron neutron has also been observed to cause transition mutations (Nakano et al., 1995). Majority of the chemical mutagens fall in the category of either intercalating agents, base modifiers, or analogs. Furthermore, these chemical agents can provide specific types of mutations for instance ethyl methane sulfonate selectively alkylates guanine residues which results in error prone replication of DNA. Hydroxylamine, hydrazine, nitrous acid, and bisulfite are also capable of causing specific transition mutations. Regardless of such specificity, the issue of inefficient control over the position and number of target residues being mutated in the genome persists for chemical mutagens (Schroeder et al., 2018). Opposed to transition and transversion mutation, insertion creates mutations that are site-specific. Such inserting elements often causes inactivation of genes and causes formation of knockout mutants. Mariner, Mu, Tn3 derivatives (Tn4430 and Tn917), Tn5, modified Tn7, Tn10, Tn552, and Tyl are well known transposable elements. Insertion site for Tn7 transposon has been identified in Gram-negative species such as *Desulfovibrio desulfuricans*, *E. coli*, *Serratia marcescens*, *Sphingomonas yanoikuyae*, *Vibrio anguillarum*, and *Vibrio fischeri* (Choi & Kim, 2009). EZ-Tn5 transposon has been utilized to create mutations within the genome of *Burkholderia seminalis* that can improve its antifungal activity against *Rhizoctonia solani* (Zhang et al., 2020). Site directed mutagenesis can create mutations that are specific as well as the position, type, and number mutations are controlled. Oligonucleotide-directed mutagenesis, cassette mutagenesis and PCR based site-directed mutagenesis has been utilized to create mutations within the genome of *E. coli*, *Brucella abortus* and *Pseudomonas aeruginosa* (Bird et al., 2012; Fasaai, 2017; Sauer et al., 2016).

Although effective afore mentioned methods of genetic manipulation are not capable of participating in genetic recombination. DNA shuffling,

in-out, and linear fragment methods utilize principles of genetic recombination for genetic manipulation. DNA shuffling is utilized to create *in vitro* recombinants of homogenous gene sequences where genes are randomly fragmented using DNaseI and after multiple rounds of PCR cycle end hybrids sequences of homologs genes are obtained (Stemmer, 1994). The in-out method utilizes multicopy plasmid that can integrate itself into the target site where crossover events take place (Madyagol et al., 2011). In linear fragment method red system, possessing three genes *game*, *bet*, and *exo* present in bacteriophage λ , plays a crucial role (Murphy et al., 2000). Clustered regularly interspaced short palindromic repeats (CRISPR) is DNA encoded defense mechanism of archaea and bacteria that has been utilized to generate controlled mutations in the genome of various microbes including *E. coli*, *S. aureus*, *Lactobacillus reuteri*, *Clostridium beijerinckii*, *Streptococcus pneumonia*, and *Saccharomyces cerevisiae*. Furthermore, advancements in the CRISPR-Cas system have resulted in the formation of cassettes using an *Acidaminococcus* sp. Cas12a-based CRISPR system that can produce multiple mutations in multiple genes simultaneously (Krappmann, 2017; Liu et al., 2019). Along with various strategies, vectors to transfer gene of interest in microbial genomes has also been of great importance. Some of the most frequently utilized vectors for microbial genetic manipulation are plasmid, bacteriophage, cosmid, bacterial artificial chromosome, and shuttle vectors. Most plasmid vectors are considered ideal for the transfer of small gene fragments because the transfer of larger DNA inserts often results in reduced stability (Rodriguez & Denhardt, 2014). Apart from plasmids bacteriophages such as λ and M13 phage also has been utilized as a vector nonetheless they possess certain limitations for instance their host range is much narrower compared to plasmids also sometimes Rec DNA causes the initiation of the lytic cycle (Nicastro et al., 2014). To overcome this problems, artificial vectors such as cosmid and bacterial artificial chromosome vectors have been created and successfully utilized in microbes (Preston, 2003).

11.4 FOOD BIOTECHNOLOGY AND USE OF GENETICALLY ENGINEERED (GE) BACTERIA

For hundreds of years, the microorganism is exploited and plays a vital role in the food production and fermented products like yogurt, cheese, wine,

vinegar, sauerkraut, etc. Multiple technologies are applied to modify the properties of microbes associated with the fermented products, increase the production of essential ingredients in food products, and improve the properties of the existing protein, elevating nutritional composition, taste and texture, aroma, and yield, thus improving the overall quality of food products. In market various food products are commercially available which is derived from the GM microbes such as, yogurt, cheese, milk, soy, food additives, amino acids, vitamins, and other nutritious food (Mallikarjuna & Yellamma, 2018). The most significant groups of bacteria in the food processing industry are lactic acid bacteria (LAB). This group includes genera such as *Lactobacillus*, *Leuconostoc*, *Lactococcus*, *Pediococcus*, *Streptococcus*, etc. Besides, various yeasts are also used in baking and brewing industries. LAB is widely used as a starter culture for the production of food products and also as a preservative. The bottlenecks such as low productivity, survival in fermentative condition, stability, are overcome with GE microbial strain. Thus, genetically improved microbial strains are widely incorporated in the place of wild strain for fermentative production of desired food products.

Application of genetic and metabolic engineering in the LAB gains significant attention as it stimulates the flavors and health beneficiary of fermented products. Diacetyl is an important flavoring compound in fermented products and also reported as GRAS (generally recognized as safe). Many microorganisms are identified for the production of diacetyl, such as, heterolactic acid bacteria, *Achromobacter lipolyticum*, *Propionibacterium*, *Aerobacter*, and *Enterobacter* species, yeast, and *Streptococci* (Zhao et al., 2009). Metabolic engineering is well exploited in many microbes for enhanced production of diacetyl and acetoin.

Guo et al. (2012) effectively carried out the partial redistribution of the glycolytic flow from lactate to diacetyl at the pyruvate branch in *Lactococcus lactis* by regulating the expression of the *noxE* gene (H_2O -forming NADH oxidase). This may be accomplished by constructing a constitutive promoter library by randomizing the gap region in the *noxE* gene's promoter. The results reported decrease in the lactate yield from 21.15 ± 0.08 mM to 9.94 ± 0.07 mM, and increase in the diacetyl concentration from 1.07 ± 0.03 mM to 4.16 ± 0.06 mM. Moreover, under aerobic conditions, the increased NADH oxidase activity from 9.43 to 58.17 folds as compared to wild strain is also notified. Thus, the promoter engineered strain eliminates the branching flux, provides better desirable end product

and depict elongated survival of bacteria cells, and improves the efficiency of starter culture in dairy fermentation industries. The presence of diacetyl and acetoin are responsible aids buttery aroma to the food product. The engineered approach applied in *Lactobacillus casei*, in which mutation in lactate dehydrogenase (Ldh) leads to the overexpression of α -acetolactate synthase (*als*) or acetohydroxyacid synthase (*ilvBN*) genes. Mutation diverts the metabolism of pyruvate to diacetyl production instead of lactate. The increased production capacity of 1,400 mg/L at pH 5.5 of diacetyl/acetoin from lactose in whey permeate is reported by the mutation in *Lactobacillus casei*, which earlier shows marginal diacetyl/acetoin production (Nadal et al., 2009). The production of diacetyl, an important ingredient in food, is increased by the knockout of α -acetolactate decarboxylase (Ald) encoding gene (*budA*) in *Enterobacter cloacae* subsp. *dissolvens* strain SDM. This causes the silencing of two diacetyl reductases (Dr) producing genes, DR-I (*gdh*) and DR-II (*budC*), which lower diacetyl synthesis substantially, either separately or in combination. The engineered *E. cloacae* SDM ($\Delta budA \Delta budC$) shows significant diacetyl production at concentration of 1.45 g/L with 0.13 g/(L·h) productivity (Zhang et al., 2015). The UV mediated mutation in wild *Enterobacter aerogenes* 10293, results in the attenuation in the genes responsible for the production of Ald, Ldh, and diacetyl reductase (Dr). The production of diacetyl is increased by 18.7 times in the UV mutated strain (Zhao et al., 2009). Pyruvate producing strain *Candida glabrata*, is metabolically engineered for diacetyl production by diverting the carbon flux to diacetyl production pathway, eliminating ILV 5 gene and BDH gene, responsible for branching pathway of α -acetolactate, and degradation of diacetyl, respectively (Gao et al., 2014).

One of the important compounds for induction of aroma is acetaldehyde (Bongers et al., 2005; Kalnenieks et al., 2019). Implementation of genetic engineering in the lactic acid bacteria (LAB) for the enhanced production of acetaldehyde induces the flavor and aroma in yogurt and other dairy products. In the food and pharmaceutical sectors, sorbitol, and mannitol are extensively used ingredients. In *Lactobacillus plantarum*, the overexpression of *stlDH* leads to increase production of sorbitol (Ladero et al., 2007). De Boeck et al. (2010) reported that the modification in the genes of *Lactobacillus casei* strain BL232 enables the increase in the sorbitol production. In *Lactobacillus reuteri* ATCC 55730, insertion of 6-phospho-1-fructokinase (*tpfkA*) gene and its activator *pkaC*

from NRRL 2270 strain of *Aspergillus niger*, leads to significant increase in the mannitol production. The engineered strain reported the production of 56 g/L mannitol, in contrast, 10 g/L is produced by parent strain. This demonstrates the effectiveness of the engineered strain in increased mannitol production (Papagianni & Legiša, 2014). Food industry shows the wide use of lactic acid as a flavoring and preservative agent. Lactic acid-tolerant yeast, *Saccharomyces cerevisiae* is constructed by the random addition of 13 genes, *HXT7*, *HXX2*, *PGII*, *PFK1*, *PFK2*, *FBA1*, *TPII*, *TDH3*, *PGK1*, *GPM1*, *ENO2*, and *PYK2*, and D-lactate dehydrogenase (D-LDH) from *Leuconostoc mesenteroides* using CRISPR-Cas system. About 33.9 g/L lactic acid is produced by the engineered strain from 100 g/L glucose, in the absence of neutralizing agent (Mitsui et al., 2020). Some of the modified organism and the desired end food product is listed in Table 11.1.

TABLE 11.1 List of Modified Organism, Targeted Genes, and Desired Product

Sl. No.	Bacterial Strain	Targeted Gene	Desired Product	References
1.	<i>Lactococcus lactis</i>	<i>noxE</i> gene	Diacetyl	Guo et al. (2012)
2.	<i>Lactobacillus casei</i>	<i>ilvBN</i> genes	Diacetyl/acetoin	Nadal et al. (2009)
3.	<i>Enterobacter cloacae</i> subsp. <i>dissolvens</i> strain SDM	<i>budA</i> gene	Diacetyl	Zhang et al. (2015)
4.	<i>Lactococcus lactis</i>	<i>dar-bdh</i> operon	homo-diacetyl and homo-(S,S)-2,3-butanediol	Liu et al. (2016)
5.	<i>Candida glabrata</i>	<i>ILV 5</i> and <i>BDH</i> gene	Diacetyl	Gao et al. (2014)
6.	<i>Lactococcus lactis</i>	<i>pdc</i> gene	Acetaldehyde	Bongers et al. (2005)
7.	<i>Streptococcus thermophilus</i>	<i>GlyA</i> gene	Acetaldehyde	Chaves et al. (2002)
8.	<i>Lactobacillus plantarum</i>	<i>stlDH</i> gene	Sorbitol	Ladero et al. (2007)
9.	<i>Lactobacillus reuteri</i> ATCC 55730	<i>tpfkA</i> and <i>pkaC</i>	Mannitol	Papagianni & Legiša (2014)
10.	<i>Lactobacillus sakei</i> B2-16	<i>GAD</i> gene	γ -aminobutyric acid	Kook et al. (2010)
11.	<i>Lactobacillus brevis</i>	<i>GAD</i> gene and <i>gad</i> operon	γ -aminobutyric acid	Wu et al. (2017)
12.	<i>Lactococcus lactis</i>	<i>folKE</i> gene	Folate	Sybesma et al. (2003)

11.5 GENETICALLY ENGINEERED (GE) MO AND REMEDIATION

11.5.1 PETROLEUM

Industries and daily life highly depend on petroleum-based products as a source of energy. During exploration, transportation, storage, other related activities, accidental leakage, and spills are common. Release of petroleum hydrocarbon may trigger adverse physical, chemical, and toxic ecological changes which affect the environmental condition. The hydrocarbon concentration and the quantity of spills are the two critical factors determining its toxicity to the biological world, either extreme or long-lasting (Ndimele, 2017). The diverse metabolic capabilities of microorganisms make them a prominent alternative to degrade and remove environmental pollutants, including petroleum products (Das & Chandran, 2011). Using advanced scientific tools and technologies, artificial microbial consortia and GMO can be created. Bacteria, fungus, algae, and cyanobacteria are some of the microorganisms that can utilize hydrocarbon as a source of nutrients and can degrade under various environmental conditions. In the laboratory condition, genetically transformed organisms can be created by utilizing molecular technologies to transfer a plasmid containing desired genes from one organism to the target host (Peixoto et al., 2011). *Acinetobacter baumannii* S30, isolated from the hydrocarbon-contaminated site, is recombinantly prepared into *Acinetobacter baumannii* S30 pJES by the insertion of the lux gene from the luciferase gene cassette luxCDABE. Recombinant *Acinetobacter baumannii* S30 pJES shows 89.3 to 53.9 g/kg soil reduction of total petroleum hydrocarbon in 90 days. However, in the shake flask condition, 50% reduction is achieved in 120 hrs. The lux inserted recombinant strain is stable in the hydrocarbon contaminated condition and can be used as an effective marker (Mishra et al., 2004). *Acinetobacter venetianus* strain RAG-1 shows excellent degradation capability, tolerance to salinity, pH, and heavy metals, and better emulsifying activity. The mutant created by the deletion of *alkMa* or *alkMb* gene demonstrate high degradation capacity for C₁₀–C₂₀ or C₂₂–C₃₂ n-alkanes, respectively. The consortia of the mutant achieved tremendous degradation efficiency of 73.42% to 88.65% for light crude oil, 68.40% to 90.05% for viscous crude oil, 47.46% to 60.52% for high waxy crude oil (Jia Liu et al., 2021). The key enzyme for aromatic hydrocarbon, Catechol 2,3-dioxygenase (*C230 gene*) are isolated from *Pseudomonas putida* strain BNF1 are introduced into NotI-cut transposon vector pUT/mini-Tn5 to produce

pUT/mini-Tn5-C23O vector. The transposon vector pUT/miniTn5-C23O are incorporated into *Acinetobacter* sp. BS3 to produce recombinant *Acinetobacter* sp. BS3-C230. The studies reported that the recombinant strain demonstrated better stability, enzyme activities and shows broad substrate specificity and thus can be used as a potential agent for the biodegradation of crude oil (Xie et al., 2014).

11.5.2 POLYCYCLIC AROMATIC HYDROCARBONS

Ubiquitously present, one of the major hazardous organic pollutants is a polycyclic aromatic hydrocarbon. Several studies have demonstrated that the GE organisms are highly capable of degrading PAH, which has opened new opportunities for research in the field of bioremediation. Physical factors of the water and soil, such as pH, temperature, water content, redox potential, solubility, contaminants, microbial populations, nutrients, organic matter, volatility, particle size, have a beneficial impact on the microbial bioremediation process (Chen et al., 2016).

GE *E. coli* is created by insertion of catechol 2,3-dioxygenase (C230) gene cloned in a suitable vector. The engineered *E. coli* shows significant decomposition of phenanthrene and pyrene in spiked soil as compared to autoclaved soil with wild *E. coli* strain and soil containing natural microflora (Ahankoub et al., 2020). Similarly, GM *Pseudomonas putida* (*P. putida*) containing catechol 2,3-dioxygenase (C230) encoded gene (*nahH*) shows remarkable removal of phenanthrene and pyrene from the soil contaminated with the spiked oil (Mardani et al., 2017). A significant n-hexane degradation capability is shown by the recombinant *Streptomyces coelicolor* M145, constructed with the overexpressing *alkB* gene encoding alkane monooxygenase (Gallo et al., 2012). The protoplast fusion between *Sphingomonas* sp. GY2B and *Pseudomonas* sp. GP3A, results into fusant strain 14 which exhibits high phenanthrene degradation capacity. The fusant strain shows better degradation efficiency and wide temperature and pH adaptation as compared to parent strains (Lu et al., 2013). Manganese peroxidase (MnP) gene, *mnp3* from *Cerrena unicolor* BBP6 is cloned and expressed in *Pichia pastoris*. The recombinant rMnP3-BBP6, shows effective fluorene and phenanthrene degradation capabilities (Zhang et al., 2020). Microbial consortia are constructed from GE *Aspergillus niger* expressing lignin peroxidase (LiP+5 strain) gene and

Phanerochaete chrysosporium expressing manganese peroxidase (MnP+7 strain) gene. The result shown that the GEM consortia exhibit high tolerance, prolong survival and enhanced degradation efficiency for PAHs (Zafra et al., 2017). Recombinant *Pseudomonas putida* KT2440-rhlABRI is prepared by the cloning of rhlABRI cassette containing rhamnolipid-producing genes obtained from *Pseudomonas aeruginosa* strain BSFD5, into the chromosome of wild *Pseudomonas putida* KT2440. Dissolution of pyrene and its degradation is elevated by the recombinant strain (Cao et al., 2012). Recombinant *Pseudomonas* sp. CGMCC2953-pK harbors catechol 2,3-dioxygenase (C23O gene) cloned into vector plasmid pK4. An increased phenanthrene biodegradation ability is presented by the recombinant strain as compared to wild strain (Zhou et al., 2013). The pyrethroid hydrolase gene (pytH) gene is extracted from *Sphingobium* sp. JZ-2, isolated from petroleum-contaminated soil, is cloned into vector pBBRMCS-5. It is then transferred into *Sphingobium* sp. BA3 to form modified *Sphingobium* sp. BA3-pytH. The recombinant strain degrades 95% of mg/L fenpropathrin in 48 h whereas, under the same condition wild strain degrade 60% of mg/L fenpropathrin (Duan et al., 2011).

Genetic engineering provides opportunities for the construction of GEM and its incorporation into the bioremediation processes. However, certain limitation bound the application to GEM merely up to laboratory scale only. Before implementing PAH treatment at the field scale, approval for releasing GEMs into the polluted site, their destiny, monitoring, control, and risk assessment for the environment are all aspects to consider.

11.5.3 DYES

A copious amount of toxic chemicals and dyes are deposited in the natural ecosystem due to rapid urbanization, causes risk to the human health and natural flora and fauna. Synthetic dyes are widely used in various industries such as textile, paper, printing, dyeing, pharmaceutical, cosmetics, and other industries. Complexity in the structure makes the synthetic dyes less susceptible to degrade in the natural condition by native microflora and traditional methods. Toxic and carcinogenic aromatic compounds are produced from synthetic dyes, which directly or indirectly shows negative impact on the ecosystem and is considered as a threat to the environment globally. Presence of dyes in the waterbodies reduce gas solubility, quality

of water and water transparency which directly or indirectly hinders the growth of natural microflora and pollute hydrosphere (Liu et al., 2019). Dyes are classified on the basis of their structure and field of application, i.e.: (i) on the basis of structure: Azo dye, nitro dye, phthalein dye, Triphenyl methane dye, indigoid dye and anthraquinone dye; (ii) on the basis of application: acid dye, basic dye, direct dye, ingrain dye, disperse dye, moderate dye, vat dye and reactive dyes (Varjani et al., 2020). Various physical, chemical, and biological methods, are successfully applied to degrade dye from the effluent (Cao et al., 2019).

Microbes-based remediation of dye contaminated wastewater is reported as more cost-effective, eco-friendly, and sustainable than traditional treatment methods (Kumar et al., 2019). Dye degradation and decolorization ability are demonstrated by various microorganisms, such as algae, bacteria, and fungi (Tochhawng et al., 2019). Along with these, physical parameters also affect dye degradation. The application of GE microorganism reveals high opportunity for the remediation processes, especially dye degradation. GEMs are prepared by the insertion of the desired gene, which plays a crucial role in the degradation of the dye, into native microbes, which overall boosts the degradation efficiency.

Sandhya (2008) reported that the recombinant *E. coli* SS125, harboring azo-reductase gene from *Bacillus latrosporus* RRK1 demonstrated decolorization efficiency of 200 mg/l Remazol Red (Azo dye) at 30°C at 255 mg cell/l/h. Whereas, no decolorization is observed with the native *E. coli* DH5α. The stability of the recombinant *E. coli* SS125 in the batch operation is also reported. Higher decolorization ability for direct blue 71 (DB 71) is reported by GE *E. coli* JM109 (pGEX-AZR). It shows the reduction of DB 71 from 150 mg/L to 27.4 mg/L in just 12 h at pH 9. However, no effect is observed at pH 5 and ranging salt from 1% to 3% (Jin et al., 2009). GE microorganism with the desired gene extracted from the source and its application for the degradation of the dyes are listed in Table 11.2.

11.5.4 HEAVY METALS

Industrialization, urbanization, and increasing population escalate the level of pollution in vital natural sources such as land and water. This rise in increasing industrialization the deposition of the toxic heavy metal in the natural environment, which readily affects human health as well as disturb

TABLE 11.2 Genetically Engineered Microorganisms Applied for the Degradation of the Dyes

Sl. No.	Genetically Engineered Microorganism	Desired Gene	Gene Extracted Source	Degradation Dye	References
1.	<i>Escherichia coli</i> JM109 (pGEXAZR)	Azoreductase gene	<i>Rhodobacter sphaeroides</i> AS1.1737	Acid red GR	Jin et al. (2008)
2.	<i>Escherichia coli</i> JM109 (pGEXAZR)	Azoreductase gene	<i>Rhodobacter sphaeroides</i> AS1.1737	Direct blue 71	Jin et al. (2009)
3.	<i>Escherichia coli</i> SS125	Azoreductase gene	<i>Bacillus latrosporus</i> RRK1	Remazol red	Sandhya et al. (2008)
4.	<i>Escherichia coli</i> CY1	Azoreductase gene	<i>Rhodococcus</i> sp.	Reactive black B	Yeh & Chang (2004)
5.	<i>Escherichia coli</i> CY1	Azoreductase gene	<i>Rhodococcus</i> sp.	Reactive red 22	Chang & Lin (2001)
6.	<i>Escherichia coli</i> DH5 α	<i>AzoA</i> gene	<i>Enterococcus</i> sp. L2	Azo dyes	Rathod et al. (2017)
7.	<i>Pseudomonas fluorescens</i> PfO-1	<i>AzoA</i> gene	<i>Enterococcus</i> sp. L2	Azo dyes	Rathod et al. (2017)
8.	<i>Escherichia coli</i> BL21 (DE3)	<i>AzoG</i> gene	<i>Halomonas</i> sp. strain GT	Azo-dye wastewater	Tian et al. (2019)
9.	<i>Escherichia coli</i> BL21	<i>AzoK</i> gene	<i>Klebsiella pneumonia</i> MGH 78578	Methyl orange	Dixit & Garg (2019)

the balance of the ecosystem (Fulekar et al., 2009). Aluminum, chromium, cadmium, cobalt, zinc, copper, manganese, nickel, mercury, and lead are considered as a global menace, as they impose negative consequences for human health and the habitat (Porter et al., 2017). Various factors are responsible for discharging heavy metal in the environmental bodies, such as mining, power stations, disposal of industrial wastes, electronic waste, rock erosion, use of pesticides and fertilizers, etc. (Vadkertiová & Sláviková, 2006). Bioavailable form of heavy metals has a tendency to enter the food web via different mechanisms. Heavy metal toxicity may cause irreversible dysfunction of the lungs, kidney, liver, brain, and nervous system, blood complications (Liu et al., 2019). Many traditional methods exist which are employed for the mitigation of the heavy metals from the contaminated site. Such conventional methods are membrane filtration, ion exchange, flocculation, adsorption, coagulation, reduction or oxidation, electrochemical treatment, reverse osmosis, evaporation, and chemical precipitation (Rajasulochana & Preethy, 2016). However, these methods are extravagant and generate a tremendous amount of secondary waste products (Bilal et al., 2017). Microbes-based bioremediation arose as a solution for the cost-intensive methods. Various microorganisms are reported to detoxify and remediate heavy metal contamination from soil and water (Jin et al., 2018).

Number of engineering strategies are applied in the development of bioremediation technologies, including genetic engineering, proteomics, and metabolic engineering, which remarks its potential in the removal of pollutants, including heavy metals (Wood et al., 2008). GE microorganism shows exclusive remediation capabilities, thus received high attention. The most exploited and implemented technology for the modification of microbial strain for heavy metal biosorption is RNA recombinant technology (Saurabh & Singh, 2017). Metal regulatory gene. Metal regulatory gene, a key factor on which the entire metal detoxification process is depended, must be present in GEM.

GM strain, *Deinococcus radiodurans* is constructed by the incorporation of *phoN*, nonspecific acid phosphatase gene, obtained from *Salmonella enterica* serovar Typhi. Bioprecipitation of 90% of the uranium from a 0.8 mM uranyl nitrate solution shown by GM strain in 6 hours (Appukuttan et al., 2006). *Deinococcus radiodurans* harboring cloned expression of *tod* and *xyl* genes of *Pseudomonas putida*, shows dual action of degradation of toluene and reduction of Cr(VI) (Brim et al., 2006). Recombinant

Caulobacter crescentus JS4022/p723–6H with RsaA–6H is fusion protein demonstrated effective sequester Cd(II). Recombinant strain and control strain shows 94.3–99.9% and 11.4–37% removal of Cd(II) respectively, within 15 min contact time (Patel et al., 2010). *merT–merP* protein and metallothionein (MT), in the recombinant *E. coli* JM109 reported potential mercury bioaccumulation in LB medium. The recombinant strain shows propagation up to the concentration of 7.4 mg/L and 96% mercury removal ability (Zhao et al., 2005). List of the heavy metal remediated by the mean of GM microorganisms are addressed in Table 11.3.

TABLE 11.3 List of the Heavy Metal Remediated by the Mean of Genetically Modified Microorganisms

Sl. No.	Genetically Engineered Microorganism	Targeted Modification	Heavy Metal Remediation	References
1.	<i>Deinococcus radiodurans</i>	<i>merA</i> gene	Mercury (Hg)	Brim et al. (2000)
2.	<i>Sphingomonas desiccabilis</i> and <i>Bacillus idriensis</i> strains	<i>arsM</i> gene	Arsenic (As)	Liu et al. (2011)
3.	<i>Pseudomonas putida</i> KT2440	<i>arsM</i> gene	Arsenic (As)	Chen et al. (2014)
4.	<i>Pseudomonas aeruginosa</i>	<i>CadR</i>	Cadmium (Cd)	Tang et al. (2018)
5.	<i>Escherichia coli</i>	<i>PcPCS1</i>	Cadmium (Cd), copper (Cu), sodium (Na), and mercury (Hg)	Li et al. (2015)
6.	<i>Methylococcus capsulatus</i>	CrR genes	Chromium Cr(VI)	Hasin et al. (2010)

11.5.5 HERBICIDES AND PESTICIDES

Pesticides have become an integral element of agricultural operations, and they are routinely utilized to fulfill growing food and feed demand. Pesticides such as organochlorines, organophosphates, carbamate, pyrethroids, chloronicotinyl, and others are used to protect plants from diseases and insects. Only a small portion of pesticides applied are used to destroy target pests, leaving the rest to contaminate cereal grains, vegetables, and fruits, or damage the environment. Pesticides are accumulated in nontarget organisms and the environment in about 99% of cases. Herbicides containing S-triazine have been identified as possible human carcinogens. Microbial degradation is a more efficient and cost-effective pesticide breakdown approach (Sehrawat et al., 2021).

The pesticides are used as nutrients by these microbes and degrade them, breaking them into little nontoxic molecules. Microbes that degrade pesticides come from a variety of microbial communities, including bacteria, fungi, actinomycetes, and algae. Pesticide-degrading bacteria include *Pseudomonas* spp., *Bacillus* spp., *Burkholderia* spp., *Klebsiella* spp., *Streptomyces*, and others, while fungi include *Trichoderma* spp., *Aspergillus* spp., *Phanerochaete chrysosporium*, white rot mushrooms, and others, while algae include *Chlamydomonas* (Sehrawat et al., 2021). Mineralization and co-metabolism are two major reactions in pesticide destruction.

The increased pollution in the environment prompted researchers to investigate microorganisms and develop genetically engineered microbes (GEMs) for contaminant removal through bioremediation. GEMs (genetically engineered microorganisms) are microbes that have been modified *in vitro* using molecular biology methodologies. As genes have been inserted into a single bacterium, GEMs contain the characteristics of numerous microorganisms. GEMs or their enzymes are directly applied to the contaminated site for breaking down the pollutants. Bioremediation has been shown to be the only viable method for decontaminating contaminated sites in a healthier, safer, long-term, and cost-effective manner (Jacob et al., 2018).

Recombinant DNA (rDNA) technology facilitates a quick and easy way to transform any organism into the desired shape by successfully incorporating gene of interest into the vector and allowing it to be expressed into a desired host (Singh & Singh, 2011).

Enzymes are the metabolites playing a vital role in the degradation of the pollutant. Pesticides may be transformed or degraded using enzymes, which is a novel treatment method for removing these chemicals from contaminated environments. Pesticide degradation catalyzed by enzymes could be more effective than current chemical processes. Hydrolases, esterases (also hydrolases), mixed function oxidases (MFO), and glutathione S-transferases (GST) are the predominant enzyme systems involved in pesticide degradation in the first metabolism stage, and the GST system in the second metabolism stage (Tahri et al., 2013). *Moraxella* sp., a naturally occurring soil organism that feeds on p-nitrophenol (PNP), has been GM to degrade both organophosphorus (OP) pesticides and PNP (Shimazu et al., 2001). *Sphingobium* sp. JQL4-5, a gram-negative fenpropathrin-degrading bacterial strain, was extracted from an insecticide factory's

wastewater treatment sludge. Other pyrethroid insecticides were degraded by strain JQL4-5, but it was unable to degrade methyl parathion. A methyl parathion hydrolase gene (*mpd*) was successfully inserted into the chromosome of strain JQL4-5 using a mini-Tn-transposon system to increase its substrate degrading range. The result was JQL4-5-*mpd*, a genetically engineered microorganism (GEM) capable of simultaneously degrading methyl parathion and fenpropathrin (Yuanfan et al., 2010). GM strain of *E. Coli* expressing the fusion protein of enhanced green fluorescent protein (EGFP) and carboxylesterase B1 (CarE B1) was successfully constructed (DE3) and successfully employed to degrade pesticides in the affected sites (Li et al., 2007). In highly atrazine-contaminated soils, GM microorganisms overexpressing catabolic genes greatly accelerate degradation. Using *E. coli* to eliminate residual atrazine contamination in soil that had been contaminated with 29 g L⁻¹ atrazine *in situ*. Recombinant *E. coli*-expressing atrazine chlorohydrolase encapsulated in organically modified silica gel was found to have superior biodegradation of atrazine (Benson et al., 2018).

11.6 GENETICALLY ENGINEERED (GE) MO AND THERAPEUTICS

GEM are becoming a promising therapeutic agent owing to the advancements in synthetic biology. Bacteria have been used to treat patients with a range of infections for more than a century. The administration of the microbiota from the healthy individual may interact with the native of the human body and respond and cure a range of diseases. Acquiring the disease's molecular and genetic information, many efficient manipulation technologies are continuously exploited to prepare the GE bacteria strain to target the particular action site and cure the disease. Moreover, the GE microorganism provides the release of complex drugs in a controlled and strategic manner (Álvarez & Fernández, 2017). GE bacteria, *in situ*, can be easily programmed and controlled using biological tools and harbor the properties of the fast, effective, and accurate detection and treatment of the disease (Lim & Song, 2019). The unique characteristics such as stress resistance, desired niche colonization, delivery of the therapeutic drug in an effective and controlled manner, and biocontainment are considered and must be present in the engineered microbial strain. Furthermore, the selection of an effective and appropriate microbial strain from the

microbiota is a key component in the creation of the designed therapeutic agent. The desired strain must be non-pathogenic, amendable for genetic manipulation, safe to the other native microbiota, and ideally survive on the site of therapeutic actions (Lim & Song, 2019).

11.6.1 BACTERIAL AND VIRAL INFECTION

Treatment of the infectious disease is as important for serving major health goals. Some of the major strategies applied for the treatment of the infectious diseases are neutralization of the toxins, hindering the adhesion of the pathogens, and interference with quorum sensing (QS) signaling (Hwang et al., 2016). Pathogenicity and virulence factors of the pathogenic microbes can be suppressing or eliminating by the use of GEM. GE *E. coli* harboring anti-biofilm enzyme gene, *DspB* (dispersion), is constructed to detect and dispense *Pseudomonas aeruginosa* under *in vitro* condition. Stability and integrity of the biofilm formation is the function of 1,6-*N*-acetyl-D-glucosamine, which is readily hydrolyzed by *DspB*. *In vivo* analysis of engineered strain on *Caenorhabditis elegans* and mice shows reduction in *P. aeruginosa* infection. Thus, the engineered bacteria known to induce prophylactic and diagnostic activity against gut infection mediated by *P. aeruginosa* (Hwang et al., 2017). *E. coli* is engineered to detect the presence of *Vibrio cholerae*, a causative agent for cholera. Engineered *E. coli* harbors CqsS, LuxU, and LuxO proteins which serves as a sensor to regulate gene expression of pQrr4, a promoter from *V. cholerae*. CRISPRi technology is used for formation of genetic circuit. The engineered *E. coli* repress the production of GFP on the detection of *Vibrio cholerae* (Holowko et al., 2016). The therapeutic engineered bacteria are highly selected for the single specific pathogen, may eliminate the excess use of antibiotics and drugs and also reduce the formation of antibiotic resistance strains (Álvarez & Fernández, 2017). Engineered *E. coli* secretes pyocin in response to the QS signal from *Pseudomonas aeruginosa*. Moreover, Engineered *E. coli* resulted in a 99% reduction of viable cells and 90% inhibition of the biofilm formation. Thus, arose as an antimicrobial strategy against the infection caused by pathogens such as *P. aeruginosa* (Saeidi et al., 2011).

11.6.2 TREATMENT OF METABOLIC DISORDER

Introduction of the appropriate engineered bacteria strain may help in the eradication of the metabolic disease's symptoms. Lack of the ability to metabolize phenylalanine (Phe) is the characteristic of the genetic disorder, phenylketonuria. Engineered *E. coli* Nissle expressing Phe-metabolizing enzyme encoding genes can be used in response to anoxic condition in mammalian guts. Around 38% blood phenylalanine reduction is reported on insertion of the engineered strain SYN1618 in mouse model. The elevation of serum phenylalanine is blood on oral Phe dietary supplementation in healthy Cynomolgus monkeys. Administration of engineered SYN1618 effectively metabolize phenylalanine into *trans*-cinnamate, which is detected in urine. Moreover, on a single oral dose, it is detected in murine and primate feces, thus acts as a predictable biomarker (Isabella et al., 2018). Glucagon-like peptide 1 (GLP-1), applied in diabetes ministration, is inserted into vector pMG36e, to form the plasmid pMG36e-GLP-1. Through electroporation, plasmid pMG36e-GLP-1 is converted into *Lactococcus lactis* MG1363 to produce GE strain MG1363-pMG36e-GLP-1 strain (M-GLP-1). The production of GLP-1 in the engineered strain greatly reduces bodyweight and increased glucose intolerance in HFD (high fat diet)-induced obese mice. Furthermore, it protects liver of obese mice by reducing the intracellular triglyceride, minimizing the level of triglycerides in the serum and enhance the expression of peroxisome proliferator-activated receptors α (PPAR α) and its target genes in the liver tissue. Thus, the engineered strain augments over all fatty acid metabolism and gut microbiome and also improves obesity in the obese mice (Wang et al., 2021). In the existence of intestinal biomarker, a nitric oxide, therapeutic drugs are released from the engineered *E. coli* which displays success in the cure of Crohn's disease (McKay et al., 2018).

11.6.3 INFLAMMATORY AND AUTOIMMUNE DISORDER

Various inflammatory and autoimmune diseases show its increasing prevalence globally. Engineered an *E. coli* strain are capable of detecting tetrathionate and also shows stability in inflammation prone mouse gut over 6 months. This engineered strain can be used as a live diagnostic agent owing to its genetic and functional durability over an extended

period of time (Riglar et al., 2017). To cure the Alzheimer's disease, GLP-1 is expressed in vector pMG36e (prokaryotic expression vector) and pLIVE (eukaryotic expression vector), which are subsequently converted into *Lactococcus lactis* subsp. *cremoris* (strain MG1363) and *Salmonella typhimurium* VNP20009 to produce GM strain MG1363-pMG36e-GLP-1 and VNP20009-pLIVE-GLP-1 strains. The strain MG1363-pMG36e-GLP-1, shows downregulation of the inflammatory responses, suppress glia activation, and prevent neuron damages. Amyloidogenesis and neuro-inflammation diseases can be cured by the application of such GE bacteria strain (Chen et al., 2018). Many engineered strains applied as a biosensor. Biosensor system by the mean of genetic engineering has been developing for the detection of tetrathionate, a biomarker for inflammatory conditions in the intestine (Riglar et al., 2017). A thiosulfate sensor was identified in *Shewanella halifaxensis* HAW-EB4, was cloned into *E. coli* Nissle 1917. Inflammation causes the activate of the biosensor in the mice (Daeffler et al., 2017).

11.7 FUTURE PROSPECTIVE

GEM, too, have pro and cons. The negative impression of the application of genetically engineered microorganisms in various is prevailed due to a lack of awareness and trust by people, authorized regulatory sectors, and scientific communities. Because of the absence of progress in the field in the course of the last decade, the utilization of GE bacteria for bioremediation has been restricted. Many precautions have been taken into consideration before releasing GEMs in the field condition. Majorly, GEMs are constructed to remediate heavy metals in laboratory conditions, ignoring the many realistic situations. The release of GEMs may harm the native microflora and disturb the ecological synchronization. Genetic engineering is depicting huge success and commendably revolutionized. Various techniques have been discovered to gain the maximum benefit of the desired gene by engineering it to the desired trait. CRISPR defeat as the revolutionary technique in genetic engineering. However, the application of GMO in food production remains contentious. The establishment of the comprehensive risk assessment processes by the regulatory authorities is mandatory for GM microorganism applied in the fermentation industries as well as food to certified it as safe and healthy as the tradition food products.

KEYWORDS

- **carbonyl cyanide m-chlorophenyl hydrazone**
- **flux balance analysis**
- **genetically engineered microorganisms**
- **genetically modified organisms**
- **horizontal gene transfer**
- **microbial cell factories**

REFERENCES

- Ahankoub, M., Mardani, G., Ghasemi-Dehkordi, P., Mehri-Ghahfarrokhi, A., Doosti, A., Jami, M. S., Allahbakhshian-Farsani, M., et al., (2020). Biodecomposition of phenanthrene and pyrene by a genetically engineered *Escherichia coli*. *Recent Patents on Biotechnology*, 14(2), 121–133. <https://doi.org/10.2174/1872208314666200128103513>.
- Álvarez, B., & Fernández, L. Á., (2017). Sustainable therapies by engineered bacteria. *Microbial Biotechnology*, 10(5), 1057–1061. <https://doi.org/10.1111/1751-7915.12778>.
- Appukkuttan, D., Rao, A. S., & Apte, S. K., (2006). Engineering of *Deinococcus radiodurans* R1 for bioprecipitation of uranium from dilute nuclear waste. *Applied and Environmental Microbiology*, 72(12), 7873–7878. <https://doi.org/10.1128/AEM.01362-06>.
- Batt, C. A., (2014). Genetic engineering. In: *Encyclopedia of Food Microbiology* (2nd edn., pp. 83–87). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-384730-0.00143-9>.
- Benson, J. J., Sakkos, J. K., Radian, A., Wackett, L. P., & Aksan, A., (2018). Enhanced biodegradation of atrazine by bacteria encapsulated in organically modified silica gels. *Journal of Colloid and Interface Science*, 510, 57–68. <https://doi.org/10.1016/j.jcis.2017.09.044>.
- Bilal, M., Asgher, M., Parra-Saldivar, R., Hu, H., Wang, W., Zhang, X., & Iqbal, H. M. N., (2017). Immobilized ligninolytic enzymes: An innovative and environmentally responsive technology to tackle dye-based industrial pollutants: A review. In: *Science of the Total Environment* (Vol. 576, pp. 646–659). Elsevier B.V. <https://doi.org/10.1016/j.scitotenv.2016.10.137>.
- Bird, A. W., Erler, A., Fu, J., Hériché, J. K., Maresca, M., Zhang, Y., Hyman, A. A., & Stewart, F., (2012). High-efficiency counterscreening recombineering for site-directed mutagenesis in bacterial artificial chromosomes. *Articles Nature Methods*, 9(1), 103. <https://doi.org/10.1038/nMeth.1803>.
- Bongers, R. S., Hoefnagel, M. H. N., & Kleerebezem, M., (2005). High-level acetaldehyde production in *Lactococcus lactis* by metabolic engineering. *Applied and Environmental Microbiology*, 71(2), 1109–1113. <https://doi.org/10.1128/AEM.71.2.1109-1113.2005>.

- Brim, H., McFarlan, S. C., Fredrickson, J. K., Minton, K. W., Zhai, M., Wackett, L. P., & Daly, M. J., (2000). Engineering *Deinococcus radiodurans* for metal remediation in radioactive mixed waste environments. *Nature Biotechnology*, 18(1), 85–90. <https://doi.org/10.1038/71986>.
- Brim, H., Osborne, J. P., Kostandarithes, H. M., Fredrickson, J. K., Wackett, L. P., & Daly, M. J., (2006). *Deinococcus radiodurans* engineered for complete toluene degradation facilitates Cr(VI) reduction. *Microbiology*, 152(8), 2469–2477. <https://doi.org/10.1099/mic.0.29009-0>.
- Burgard, A. P., Pharkya, P., & Maranas, C. D., (2003). OptKnock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnology and Bioengineering*, 84(6), 647–657. <https://doi.org/10.1002/bit.10803>.
- Cao, J., Sanganyado, E., Liu, W., Zhang, W., & Liu, Y., (2019). Decolorization and detoxification of direct blue 2B by indigenous bacterial consortium. *Journal of Environmental Management*, 242, 229–237. <https://doi.org/10.1016/j.jenvman.2019.04.067>.
- Cao, L., Wang, Q., Zhang, J., Li, C., Yan, X., Lou, X., Xia, Y., Hong, Q., & Li, S., (2012). Construction of a stable genetically engineered rhamnolipid-producing microorganism for remediation of pyrene-contaminated soil. *World Journal of Microbiology and Biotechnology*, 28(9), 2783–2790. <https://doi.org/10.1007/s11274-012-1088-0>.
- Chang, J. S., & Lin, C. Y., (2001). Decolorization kinetics of a recombinant *Escherichia coli* strain harboring azo-dye-decolorizing determinants from *Rhodococcus* sp. *Biotechnology Letters*, 23(8), 631–636. <https://doi.org/10.1023/A:1010306114286>.
- Chaves, A. C. S. D., Fernandez, M., Lerayer, A. L. S., Mierau, I., Kleerebezem, M., & Hugenholtz, J., (2002). Metabolic engineering of acetaldehyde production by *Streptococcus thermophilus*. *Applied and Environmental Microbiology*, 68(11), 5656–5662. <https://doi.org/10.1128/AEM.68.11.5656-5662.2002>.
- Chen, D., Wang, H., & Yang, K., (2016). Effective biodegradation of nitrate, Cr(VI) and p-fluoronitrobenzene by a novel three dimensional bioelectrochemical system. *Bioresource Technology*, 203, 370–373. <https://doi.org/10.1016/j.biortech.2015.12.059>.
- Chen, J., Sun, G. X., Wang, X. X., Lorenzo, V. De., Rosen, B. P., & Zhu, Y. G., (2014). Volatilization of arsenic from polluted soil by *Pseudomonas putida* engineered for expression of the arsM arsenic(III) S-adenosine methyltransferase gene. *Environmental Science and Technology*, 48(17), 10337–10344. <https://doi.org/10.1021/es502230b>.
- Chen, T., Tian, P., Huang, Z., Zhao, X., Wang, H., Xia, C., Wang, L., & Wei, H., (2018). Engineered commensal bacteria prevent systemic inflammation-induced memory impairment and amyloidogenesis via producing GLP-1. *Applied Microbiology and Biotechnology*, 102(17), 7565–7575. <https://doi.org/10.1007/s00253-018-9155-6>.
- Choi, K. H., & Kim, J., (2009). Applications of transposon-based gene delivery system in bacteria. *J. Microbiol. Biotechnol*, 19(3), 217–228. <https://doi.org/10.4014/jmb.0811.669>.
- Chowdhury, A., Zomorodi, A. R., & Maranas, C. D., (2014). k-OptForce: Integrating kinetics with flux balance analysis for strain design. *PLoS Computational Biology*, 10(2). <https://doi.org/10.1371/journal.pcbi.1003487>.
- Daeffler, K. N., Galley, J. D., Sheth, R. U., Ortiz-Velez, L. C., Bibb, C. O., Shroyer, N. F., Britton, R. A., & Tabor, J. J., (2017). Engineering bacterial thiosulfate and tetrathionate sensors for detecting gut inflammation. *Molecular Systems Biology*, 13(4), 923. <https://doi.org/10.15252/msb.20167416>.

- Das, N., & Chandran, P., (2011). *Research Biotechnology Research International*, 2011, 13. <https://doi.org/10.4061/2011/941810>.
- Davis, T. O., Henderson, I., Brehm, J. K., & Minton, N. P., (2000). Development of a transformation and gene reporter system for group II, non-proteolytic clostridium botulinum type B strains fermentation symposium JMMB research article. In: *J. Mol. Microbiol. Biotechnol* (Vol. 2, Issue 1). www.caister.com/bacteria-plant (accessed on 12 October 2022).
- Davison, J., (1999). Genetic exchange between bacteria in the environment. In: *Plasmid* (Vol. 42, No. 2, pp. 73–91). <https://doi.org/10.1006/plas.1999.1421>.
- De Boeck, R., Sarmiento-Rubiano, L. A., Nadal, I., Monedero, V., Pérez-Martínez, G., & Yebra, M. J., (2010). Sorbitol production from lactose by engineered *Lactobacillus casei* deficient in sorbitol transport system and mannitol-1-phosphate dehydrogenase. *Applied Microbiology and Biotechnology*, 85(6), 1915–1922. <https://doi.org/10.1007/s00253-009-2260-9>.
- Dixit, S., & Garg, S., (2019). Development of an efficient recombinant bacterium and its application in the degradation of environmentally hazardous azo dyes. *International Journal of Environmental Science and Technology*, 16(11), 7137–7146. <https://doi.org/10.1007/s13762-018-2054-7>.
- Duan, X. Q., Zheng, J. W., Zhang, J., Hang, B. J., He, J., & Li, S. P., (2011). Characteristics of a 3-phenoxybenzoic acid degrading-dacterium and the construction of a engineering bacterium. *Huanjing Kexue/Environmental Science*, 32(1), 240–246. <https://europemc.org/article/med/21404693> (accessed on 12 October 2022).
- Fasaei, N., (2017). Site-directed mutagenesis in *Brucella abortus* S19 by overlap extension PCR-based procedure. *Journal of the Hellenic Veterinary Medical Society*, 68(3), 273–278. <https://doi.org/10.12681/jhvms.15468>.
- Fisher, A. K., Freedman, B. G., Bevan, D. R., & Senger, R. S., (2014). A review of metabolic and enzymatic engineering strategies for designing and optimizing performance of microbial cell factories. In: *Computational and Structural Biotechnology Journal* (Vol. 11, No. 18, pp. 91–99). <https://doi.org/10.1016/j.csbj.2014.08.010>.
- Fulekar, M. H., Singh, A., & Bhaduri, A. M., (2009). Genetic engineering strategies for enhancing phytoremediation of heavy metals. *African Journal of Biotechnology*, 8(4), 529–535. <http://www.academicjournals.org/AJB> (accessed on 12 October 2022).
- Gallo, G., Lo Piccolo, L., Renzone, G., La Rosa, R., Scaloni, A., Quatrini, P., & Puglia, A. M., (2012). Differential proteomic analysis of an engineered *Streptomyces coelicolor* strain reveals metabolic pathways supporting growth on n-hexadecane. *Applied Microbiology and Biotechnology*, 94(5), 1289–1301. <https://doi.org/10.1007/s00253-012-4046-8>.
- Gao, X., Xu, N., Li, S., & Liu, L., (2014). Metabolic engineering of candida glabrata for diacetyl production. *PLoS One*, 9(3), e89854. <https://doi.org/10.1371/journal.pone.0089854>.
- Guo, T., Kong, J., Zhang, L., Zhang, C., & Hu, S., (2012). Fine-tuning of the lactate and diacetyl production through promoter engineering in *Lactococcus lactis*. *PLoS One*, 7(4), e36296. <https://doi.org/10.1371/journal.pone.0036296>.
- Gupta, S., & Singh, D., (2017). Role of genetically modified microorganisms in heavy metal bioremediation. In: *Advances in Environmental Biotechnology* (pp. 197–214). Springer Singapore. https://doi.org/10.1007/978-981-10-4041-2_12.

- Gupta, S., Pathak, B., & Fulekar, M. H., (2015). Molecular approaches for biodegradation of polycyclic aromatic hydrocarbon compounds: A review. In *Reviews in Environmental Science and Biotechnology* (Vol. 14, No. 2, pp. 241–269). <https://doi.org/10.1007/s11157-014-9353-3>.
- Han, L., (2004). Genetically modified microorganisms. In: *The GMO Handbook* (pp. 29–51). Humana Press. https://doi.org/10.1007/978-1-59259-801-4_2.
- Hasin, A. A. L., Gurman, S. J., Murphy, L. M., Perry, A., Smith, T. J., & Gardiner, P. H. E., (2010). Remediation of chromium(VI) by a methane-oxidizing bacterium. *Environmental Science and Technology*, 44(1), 400–405. <https://doi.org/10.1021/es901723c>.
- Holowko, M. B., Wang, H., Jayaraman, P., & Poh, C. L., (2016). Biosensing vibrio cholerae with genetically engineered *Escherichia coli*. *ACS Synthetic Biology*, 5(11), 1275–1283. <https://doi.org/10.1021/acssynbio.6b00079>.
- Hopwood, D. A., (1981). Genetic studies with bacterial protoplasts. In: *Annual Review of Microbiology* (Vol. 35, pp. 237–272). <https://doi.org/10.1146/annurev.mi.35.100181.001321>.
- Hwang, I. Y., Koh, E., Kim, H. R., Yew, W. S., & Chang, M. W., (2016). Reprogrammable microbial cell-based therapeutics against antibiotic-resistant bacteria. In: *Drug Resistance Updates* (Vol. 27, pp. 59–71). <https://doi.org/10.1016/j.drug.2016.06.002>.
- Hwang, I. Y., Koh, E., Wong, A., March, J. C., Bentley, W. E., Lee, Y. S., & Chang, M. W., (2017). Engineered probiotic *Escherichia coli* can eliminate and prevent *Pseudomonas aeruginosa* gut infection in animal models. *Nature Communications*, 8(1), 1–11. <https://doi.org/10.1038/ncomms15028>.
- Isabella, V. M., Ha, B. N., Castillo, M. J., Lubkowitz, D. J., Rowe, S. E., Millet, Y. A., Anderson, C. L., et al., (2018). Development of a synthetic live bacterial therapeutic for the human metabolic disease phenylketonuria. *Nature Biotechnology*, 36(9), 857–867. <https://doi.org/10.1038/nbt.4222>.
- Jacob, J. M., Karthik, C., Saratale, R. G., Kumar, S. S., Prabakar, D., Kadirvelu, K., & Pugazhendhi, A., (2018). Biological approaches to tackle heavy metal pollution: A survey of literature. In: *Journal of Environmental Management* (Vol. 217, pp. 56–70). Academic Press. <https://doi.org/10.1016/j.jenvman.2018.03.077>.
- Jin, R. F., Zhou, J. T., Zhang, A. L., & Wang, J., (2008). Bioaugmentation of the decolorization rate of acid red GR by genetically engineered microorganism *Escherichia coli* JM109 (pGEX-AZR). *World Journal of Microbiology and Biotechnology*, 24(1), 23–29. <https://doi.org/10.1007/s11274-007-9433-4>.
- Jin, R., Yang, H., Zhang, A., Wang, J., & Liu, G., (2009). Bioaugmentation on decolorization of C.I. direct blue 71 by using genetically engineered strain *Escherichia coli* JM109 (pGEX-AZR). *Journal of Hazardous Materials*, 163(2, 3), 1123–1128. <https://doi.org/10.1016/j.jhazmat.2008.07.067>.
- Jin, Y., Luan, Y., Ning, Y., & Wang, L., (2018). *Effects and Mechanisms of Microbial Remediation of Heavy Metals in Soil: A Critical Review*. Mdpi.com. <https://doi.org/10.3390/app8081336>.
- Johnsborg, O., Eldholm, V., & Sigve, H. L., (2007). *Natural Genetic Transformation: Prevalence, Mechanisms and Function*. Elsevier. <https://doi.org/10.1016/j.resmic.2007.09.004>.
- Kalnenieks, U., Balodite, E., Strähler, S., Strazdina, I., Rex, J., Pentjuss, A., Fuchino, K., et al., (2019). Improvement of acetaldehyde production in *Zymomonas mobilis* by

- engineering of its aerobic metabolism. *Frontiers in Microbiology*, 10, 2533. <https://doi.org/10.3389/fmicb.2019.02533>.
- Kim, J., & Reed, J. L., (2010). OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Systems Biology*, 4. <https://doi.org/10.1186/1752-0509-4-53>.
- Kook, M. C., Seo, M. J., Cheigh, C. I., Lee, S. J., Pyun, Y. R., & Park, H., (2010). Enhancement of γ -aminobutyric acid production by *Lactobacillus sakei* B2-16 expressing glutamate decarboxylase from *Lactobacillus plantarum* ATCC 14917. *J. Korean Soc. Appl. Biol. Chem*, 53(6), 816–820. <https://doi.org/10.3839/jksabc.2010.123>.
- Krappmann, S., (2017). CRISPR-Cas9, the new kid on the block of fungal molecular biology. *Medical Mycology*, 55(1), 16–23. <https://doi.org/10.1093/mmy/myw097>.
- Ladero, V., Ramos, A., Wiersma, A., Goffin, P., Schanck, A., Kleerebezem, M., Hugenholtz, J., et al., (2007). High-level production of the low-calorie sugar sorbitol by *Lactobacillus plantarum* through metabolic engineering. *Applied and Environmental Microbiology*, 73(6), 1864–1872. <https://doi.org/10.1128/AEM.02304-06>.
- Li, G. W., Burkhardt, D., Gross, C., & Weissman, J. S., (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell*, 157(3), 624–635. <https://doi.org/10.1016/j.cell.2014.02.033>.
- Li, H., Cong, Y., Lin, J., & Chang, Y., (2015). Enhanced tolerance and accumulation of heavy metal ions by engineered *Escherichia coli* expressing *Pyrus calleryana* phytochelatin synthase. *Journal of Basic Microbiology*, 55(3), 398–405. <https://doi.org/10.1002/jobm.201300670>.
- Li, Q., Chen, R., Li, W., Qiao, C. L., & Wu, Y. J., (2007). A genetically engineered *Escherichia coli*, expressing the fusion protein of green fluorescent protein and carboxylesterase B1, can be easily detected in the environment following degradation of pesticide residues. *Biotechnology Letters*, 29(9), 1357–1362. <https://doi.org/10.1007/s10529-007-9410-x>.
- Lim, D., & Song, M., (2019). Development of bacteria as diagnostics and therapeutics by genetic engineering. *Journal of Microbiology*, 57(8), 637–643. <https://doi.org/10.1007/s12275-019-9105-8>.
- Liu, J., Chan, S. H. J., Brock-Nannestad, T., Chen, J., Lee, S. Y., Solem, C., & Jensen, P. R., (2016). Combining metabolic engineering and biocompatible chemistry for high-yield production of homo-diacetyl and homo-(S,S)-2,3-butanediol. *Metabolic Engineering*, 36, 57–67. <https://doi.org/10.1016/j.ymben.2016.02.008>.
- Liu, J., Zhao, B., Lan, Y., & Ma, T., (2021). Enhanced degradation of different crude oils by defined engineered consortia of *Acinetobacter venetianus* RAG-1 mutants based on their alkane metabolism. *Bioresource Technology*, 327, 124787. <https://doi.org/10.1016/j.biortech.2021.124787>.
- Liu, L., Bilal, M., Duan, X., & Iqbal, H. M. N., (2019). Mitigation of environmental pollution by genetically engineered bacteria — Current challenges and future perspectives. In: *Science of the Total Environment* (Vol. 667, pp. 444–454). Elsevier B.V. <https://doi.org/10.1016/j.scitotenv.2019.02.390>.
- Liu, Q., Zhang, Y., Li, F., Li, J., Sun, W., & Tian, C., (2019). Upgrading of efficient and scalable CRISPR-Cas-mediated technology for genetic engineering in thermophilic fungus *Myceliophthora thermophila*. *Biotechnology for Biofuels*, 12(1). <https://doi.org/10.1186/s13068-019-1637-y>.

- Liu, S., Zhang, F., Chen, J., & Sun, G., (2011). Arsenic removal from contaminated soil via volatilization by genetically engineered bacteria under laboratory conditions. *Journal of Environmental Sciences*, 23(9), 1544–1550. [https://doi.org/10.1016/S1001-0742\(10\)60570-0](https://doi.org/10.1016/S1001-0742(10)60570-0).
- Lu, J., Guo, C., Li, J., Zhang, H., Lu, G., Dang, Z., & Wu, R., (2013). A fusant of *Sphingomonas* sp. GY2B and *Pseudomonas* sp. GP3A with high capacity of degrading phenanthrene. *World Journal of Microbiology and Biotechnology*, 29(9), 1685–1694. <https://doi.org/10.1007/s11274-013-1331-3>.
- Madyagol, M., Al-Alami, H., Levarski, Z., Drahovská, H., Turňa, J., & Stuchlík, S., (2011). Gene replacement techniques for *Escherichia coli* genome modification. In: *Folia Microbiologica* (Vol. 56, No. 3, pp. 253–263). <https://doi.org/10.1007/s12223-011-0035-z>.
- Mallikarjuna, N., & Yellamma, K., (2018). Genetic and metabolic engineering of microorganisms for the production of various food products. In: *Recent Developments in Applied Microbiology and Biochemistry* (pp. 167–182). Elsevier. <https://doi.org/10.1016/B978-0-12-816328-3.00013-1>.
- Mardani, G., Mahvi, A. H., Hashemzadeh-Chaleshtori, M., Naseri, S., Dehghani, M. H., & Ghasemi-Dehkordi, P., (2017). Application of genetically engineered dioxygenase producing *pseudomonas putida* on decomposition of oil from spiked soil. *Jundishapur Journal of Natural Pharmaceutical Products*, 12(3 (Supp)). <https://doi.org/10.5812/jjnpp.64313>.
- McAnulty, M. J., Yen, J. Y., Freedman, B. G., & Senger, R. S., (2012). Genome-scale modeling using flux ratio constraints to enable metabolic engineering of clostridial metabolism in silico. *BMC Systems Biology*, 6. <https://doi.org/10.1186/1752-0509-6-42>.
- McKay, R., Ghodasra, M., Schardt, J., Quan, D., Pottash, A. E., Shang, W., Jay, S. M., et al., (2018). A platform of genetically engineered bacteria as vehicles for localized delivery of therapeutics: Toward applications for Crohn's disease. *Bioengineering & Translational Medicine*, 3(3), 209–221. <https://doi.org/10.1002/btm2.10113>.
- Mishra, S., Sarma, P. M., & Lal, B., (2004). Crude oil degradation efficiency of a recombinant *Acinetobacter baumannii* strain and its survival in crude oil-contaminated soil microcosm. *FEMS Microbiology Letters*, 235(2), 323–331. <https://doi.org/10.1016/j.femsle.2004.05.002>.
- Mitsui, R., Yamada, R., Matsumoto, T., Yoshihara, S., Tokumoto, H., & Ogino, H., (2020). Construction of lactic acid-tolerant *Saccharomyces cerevisiae* by using CRISPR-Cas-mediated genome evolution for efficient d-lactic acid production. *Applied Microbiology and Biotechnology*, 104(21), 9147–9158. <https://doi.org/10.1007/s00253-020-10906-3>.
- Murphy, K. C., Campellone, K. G., & Poteete, A. R., (2000). PCR-mediated gene replacement in *Escherichia coli*. In: *Gene* (Vol. 246). www.elsevier.com/locate/gene (accessed on 12 October 2022).
- Na, D., Yoo, S. M., Chung, H., Park, H., Park, J. H., & Lee, S. Y., (2013). Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nature Biotechnology*, 31(2), 170–174. <https://doi.org/10.1038/nbt.2461>.
- Nadal, I., Rico, J., Pérez-Martínez, G., Yebra, M. J., & Monedero, V. (2009). Diacetyl and acetoin production from whey permeate using engineered *Lactobacillus casei*. *Academic. Oup. Com*. <https://academic.oup.com/jimb/article-abstract/36/9/1233/5993634> (accessed on 12 October 2022).

- Nakano, T., Okaichi, K., Harada, K., Matsumoto, H., Kimura, R., Yamamoto, K., Akasaka, S., & Ohnishi, T., (1995). Mutations of a shuttle vector plasmid, pZ189, in *Escherichia coli* induced by boron neutron captured beam (BNCB) containing α -particles. *Mutation Research-DNA Repair*, 336(2), 153–159. [https://doi.org/10.1016/0921-8777\(94\)00053-9](https://doi.org/10.1016/0921-8777(94)00053-9).
- Ndimele, P. E., (2017). The political ecology of oil and gas activities in the Nigerian aquatic ecosystem. In *The Political Ecology of Oil and Gas Activities in the Nigerian Aquatic Ecosystem*. <https://doi.org/10.1016/c2015-0-05649-0>.
- Nicastro, J., Sheldon, K., & Slavcev, R. A., (2014). Bacteriophage lambda display systems: Developments and applications. In: *Applied Microbiology and Biotechnology* (Vol. 98, No. 7, pp. 2853–2866). Springer Verlag. <https://doi.org/10.1007/s00253-014-5521-1>.
- Panja, S., Saha, S., Jana, B., & Basu, T., (2006). Role of membrane potential on artificial transformation of *E. coli* with plasmid DNA. *Journal of Biotechnology*, 127, 14–20. <https://doi.org/10.1016/j.jbiotec.2006.06.008>.
- Papagianni, M., & Legiša, M., (2014). Increased mannitol production in *Lactobacillus reuteri* ATCC 55730 production strain with a modified 6-phosphofructo-1-kinase. *Journal of Biotechnology*, 181, 20–26. <https://doi.org/10.1016/j.jbiotec.2014.04.007>.
- Patel, J., Zhang, Q., McKay, R. M. L., Vincent, R., & Xu, Z., (2010). Genetic engineering of *Caulobacter crescentus* for removal of cadmium from water. *Applied Biochemistry and Biotechnology*, 160(1), 232–243. <https://doi.org/10.1007/s12010-009-8540-0>.
- Peixoto, R. S., Vermelho, A. B., & Rosado, A. S., (2011). Petroleum-degrading enzymes: Bioremediation and new prospects. In: *Enzyme Research* (Vol. 2011, No. 1). <https://doi.org/10.4061/2011/475193>.
- Peter, R., Mojca, J., & Primož, P., (2011). Genetically modified organisms (GMOs). In: *Encyclopedia of Environmental Health* (pp. 879–888). Elsevier Inc. <https://doi.org/10.1016/B978-0-444-52272-6.00481-5>.
- Peter, R., Mojca, J., & Primož, P., (2019). Genetically modified organisms (GMOs). In *Encyclopedia of Environmental Health* (pp. 199–207). Elsevier. <https://doi.org/10.1016/B978-0-444-63951-6.00481-2>.
- Porter, S. S., Chang, P. L., Conow, C. A., Dunham, J. P., & Friesen, M. L., (2017). Association mapping reveals novel serpentine adaptation gene clusters in a population of symbiotic *Mesorhizobium*. *ISME Journal*, 11(1), 248–262. <https://doi.org/10.1038/ismej.2016.88>.
- Prakash, D., Verma, S., Bhatia, R., & Tiwary, B. N., (2011). Risks and precautions of genetically modified organisms. *International Scholarly Research Network ISRN Ecology*, 2011, 13. <https://doi.org/10.5402/2011/369573>.
- Preston, A., (2003). Choosing a cloning vector. *Methods in Molecular Biology* (Clifton, N.J.), 235, 19–26. <https://doi.org/10.1385/1-59259-409-3:19>.
- Rajasulochana, P., & Preethy, V., (2016). Comparison on efficiency of various techniques in treatment of waste and sewage water – a comprehensive review. *Resource-Efficient Technologies*, 2(4), 175–184. <https://doi.org/10.1016/j.refit.2016.09.004>.
- Ranganathan, S., Suthers, P. F., & Maranas, C. D., (2010). OptForce: An optimization procedure for identifying all genetic manipulations leading to targeted overproductions. *PLoS Computational Biology*, 6(4). <https://doi.org/10.1371/journal.pcbi.1000744>.
- Rathod, J., Dhebar, S., & Archana, G., (2017). Efficient approach to enhance whole cell azo dye decolorization by heterologous overexpression of *Enterococcus* sp. L2 azoreductase (azoA) and *Mycobacterium vaccae* formate dehydrogenase (fdh) in different bacterial

- systems. *International Biodeterioration and Biodegradation*, 124, 91–100. <https://doi.org/10.1016/j.ibiod.2017.04.023>.
- Riglar, D. T., Giessen, T. W., Baym, M., Kerns, S. J., Niederhuber, M. J., Bronson, R. T., Kotula, J. W., et al., (2017). Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nature Biotechnology*, 35(7), 653–658. <https://doi.org/10.1038/nbt.3879>.
- Rivera, A. L., Magaña-Ortiz, D., Gómez-Lim, M., Fernández, F., & Loske, A. M., (2014). *ScienceDirect Physical Methods for Genetic Transformation of Fungi and Yeast*. Elsevier. <https://doi.org/10.1016/j.plrev.2014.01.007>.
- Rodriguez, R., & Denhardt, D., (2014). *Vectors: A Survey of Molecular Cloning Vectors and Their Uses*. <https://books.google.com/books?hl=en&lr=&id=PG7iBQAAQBAJ&oi=fnd&pg=PP1&dq=Vectors:+a+survey+of+molecular+cloning+vectors+and+their+uses:+Butterworth-Heinemann.&ots=U06-DzNY0l&sig=3yUuVKKeu0WFg3TVzU6VYWDtUJ8> (accessed on 12 October 2022).
- Saeidi, N., Wong, C. K., Lo, T. M., Nguyen, H. X., Ling, H., Leong, S. S. J., Poh, C. L., & Chang, M. W., (2011). Engineering microbes to sense and eradicate *Pseudomonas aeruginosa*, a human pathogen. *Molecular Systems Biology*, 7. <https://doi.org/10.1038/msb.2011.55>.
- Sandhya, S., Sarayu, K., Uma, B., & Swaminathan, K., (2008). Decolorizing kinetics of a recombinant *Escherichia coli* SS125 strain harboring azoreductase gene from *Bacillus latrosporus* RRK1. *Bioresource Technology*, 99(7), 2187–2191. <https://doi.org/10.1016/j.biortech.2007.05.027>.
- Sauer, N. J., Mozoruk, J., Miller, R. B., Warburg, Z. J., Walker, K. A., Beetham, P. R., Sch€ Opke, C. R., & Gocal, G. F. W., (2016). *Oligonucleotide-Directed Mutagenesis for Precision Gene Editing*, 14(2), 496–502. Wiley Online Library. <https://doi.org/10.1111/pbi.12496>.
- Schroeder, J. W., Yeesin, P., Simmons, L. A., & Wang, J. D., (2018). Sources of spontaneous mutagenesis in bacteria. In: *Critical Reviews in Biochemistry and Molecular Biology* (Vol. 53, No. 1, pp. 29–48). Taylor and Francis Ltd. <https://doi.org/10.1080/10409238.2017.1394262>.
- Sehrawat, A., Phour, M., Kumar, R., & Sindhu, S. S., (2021). *Bioremediation of Pesticides: An Eco-Friendly Approach for Environment Sustainability* (pp. 23–84). https://doi.org/10.1007/978-981-15-7447-4_2.
- Senthil, K. P., Joshiba, G. J., Femina, C. C., Varshini, P., Priyadharshini, S., Karthick, A., & Jothirani, R., (2019). *A Critical Review on Recent Developments in the Low-Cost Adsorption of Dyes from Wastewater*. <https://doi.org/10.5004/dwt.2019.24613>.
- Shibai, A., Takahashi, Y., Ishizawa, Y., Motooka, D., Nakamura, S., Ying, B. W., & Tsuru, S., (2017). Mutation accumulation under UV radiation in *Escherichia coli*. *Scientific Reports*, 7(1). <https://doi.org/10.1038/s41598-017-15008-1>.
- Shimazu, M., Mulchandani, A., & Chen, W., (2001). Simultaneous degradation of organophosphorus pesticides and p-nitrophenol by a genetically engineered *Moraxella* sp. with surface-expressed organophosphorus hydrolase. *Biotechnology and Bioengineering*, 76(4), 318–324. <https://doi.org/10.1002/bit.10095>.
- Shirshikova, T. V., Morozova, O. V., Kamaletdinova, L. K., Sharipova, M. R., & Bogomolnaya, L. M., (2016). Generalized bacteriophage transduction in *Serratia marcescens*. *BioNanoScience*, 6(4), 487–489. <https://doi.org/10.1007/s12668-016-0268-z>.

- Singh, N. S., & Singh, D. K., (2011). Biodegradation of endosulfan and endosulfan sulfate by *Achromobacter xylosoxidans* strain C8B in broth medium. *Biodegradation*, 22(5), 845–857. <https://doi.org/10.1007/s10532-010-9442-0>.
- Son, Y. E., & Park, H. S., (2020). *Genetic Manipulation and Transformation Methods for Aspergillus spp.* *Genetic Manipulation and Transformation Methods for Aspergillus spp.* Taylor & Francis. <https://doi.org/10.1080/12298093.2020.1838115>.
- Stemmer, W. P. C., (1994). DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. In: *Genetics* (Vol. 91). <https://www.pnas.org/content/91/22/10747.short> (accessed on 12 October 2022).
- Stewart, G. G., & Stewart, G. G., (2017). Yeast genetic manipulation. In *Brewing and Distilling Yeasts* (pp. 357–387). Springer International Publishing. https://doi.org/10.1007/978-3-319-69126-8_16.
- Sybesma, W., Starrenburg, M., Kleerebezem, M., Mierau, I., De Vos, W. M., & Hugenholtz, J., (2003). Increased production of folate by metabolic engineering of *Lactococcus lactis*. *Applied and Environmental Microbiology*, 69(6), 3069–3076. <https://doi.org/10.1128/AEM.69.6.3069-3076.2003>.
- Tahri, N., Bahafid, W., Sayel, H., & El Ghachtouli, N., (2013). Biodegradation: Involved microorganisms and genetically engineered microorganisms. In: *Biodegradation - Life of Science*. <https://doi.org/10.5772/56194>.
- Tang, X., Zeng, G., Fan, C., Zhou, M., Tang, L., Zhu, J., Wan, J., et al., (2018). Chromosomal expression of CadR on *Pseudomonas aeruginosa* for the removal of Cd(II) from aqueous solutions. *Science of the Total Environment*, 636, 1355–1361. <https://doi.org/10.1016/j.scitotenv.2018.04.229>.
- Tian, F., Guo, G., Zhang, C., Yang, F., Hu, Z., Liu, C., & Wang, S. W., (2019). Isolation, cloning and characterization of an azoreductase and the effect of salinity on its expression in a halophilic bacterium. *International Journal of Biological Macromolecules*, 123, 1062–1069. <https://doi.org/10.1016/j.ijbiomac.2018.11.175>.
- Tochhawng, L., Mishra, V. K., Passari, A. K., & Singh, B. P., (2019). *Endophytic Fungi: Role in Dye Decolorization* (pp. 1–15). Springer, Cham. https://doi.org/10.1007/978-3-030-03589-1_1.
- Vadkertiová, R., & Sláviková, E., (2006). Metal tolerance of yeasts isolated from water, soil and plant environments. *Journal of Basic Microbiology*, 46(2), 145–152. <https://doi.org/10.1002/jobm.200510609>.
- Varjani, S., Rakholiya, P., Ng, Y., You, S., & Teixeira, J. A., (2020). *Microbial Degradation of Dyes: An Overview*. <https://doi.org/10.1016/j.biortech.2020.123728>.
- Wang, L., Chen, T., Wang, H., Wu, X., Cao, Q., Wen, K., Deng, K. Y., & Xin, H., (2021). Engineered bacteria of MG1363-pMG36e-GLP-1 attenuated obesity-induced by high fat diet in mice. *Frontiers in Cellular and Infection Microbiology*, 11. <https://doi.org/10.3389/fcimb.2021.595575>.
- Wood, T. K., Zhao, H., & Chen, W., (2008). Author's personal copy molecular approaches in bioremediation this review comes from a themed issue on chemical biotechnology edited. *Current Opinion in Biotechnology*, 19, 572–578. <https://doi.org/10.1016/j.copbio.2008.10.003>.
- Wu, Q., & Shah, N. P. (2017). High γ -aminobutyric acid production from lactic acid bacteria: emphasis on *Lactobacillus brevis* as a functional dairy starter. *Critical Reviews in Food Science and Nutrition*, 57(17), 3661–3672.

- Xie, Y., Yu, F., Wang, Q., Gu, X., & Chen, W., (2014). Cloning of catechol 2,3-dioxygenase gene and construction of a stable genetically engineered strain for degrading crude oil. *Indian Journal of Microbiology*, 54(1), 59–64. <https://doi.org/10.1007/s12088-013-0411-2>.
- Yeh, M. S., & Chang, J. S., (2004). Bacterial decolorization of an azo dye with a natural isolate of *Pseudomonas luteola* and genetically modified *Escherichia coli*. *Journal of Chemical Technology & Biotechnology*, 79(12), 1354–1360. <https://doi.org/10.1002/jctb.1099>.
- Yen, J. Y., Nazem-Bokaei, H., Freedman, B. G., Athamneh, A. I. M., & Senger, R. S., (2013). Deriving metabolic engineering strategies from genome-scale modeling with flux ratio constraints. *Biotechnology Journal*, 8(5), 581–594. <https://doi.org/10.1002/biot.201200234>.
- Yuanfan, H., Jin, Z., Qing, H., Qian, W., Jiandong, J., & Shunpeng, L., (2010). Characterization of a fenpropathrin-degrading strain and construction of a genetically engineered microorganism for simultaneous degradation of methyl parathion and fenpropathrin. *Journal of Environmental Management*, 91(11), 2295–2300. <https://doi.org/10.1016/j.jenvman.2010.06.010>.
- Zafra, G., Absalón, Á. E., Anducho-Reyes, M. Á., Fernandez, F. J., & Cortés-Espinosa, D. V., (2017). Construction of PAH-degrading mixed microbial consortia by induced selection in soil. *Chemosphere*, 172, 120–126. <https://doi.org/10.1016/j.chemosphere.2016.12.038>.
- Zhang, H., Zhang, X., & Geng, A., (2020). Expression of a novel manganese peroxidase from *Cerrena unicolor* BBP6 in *Pichia pastoris* and its application in dye decolorization and PAH degradation. *Biochemical Engineering Journal*, 153, 107402. <https://doi.org/10.1016/j.bej.2019.107402>.
- Zhang, L., Zhang, Y., Liu, Q., Meng, L., Hu, M., Lv, M., Li, K., et al., (2015). Production of diacetyl by metabolically engineered *Enterobacter cloacae*. *Scientific Reports*, 5(1), 1–7. <https://doi.org/10.1038/srep09033>.
- Zhang, M., Wang, X., Ahmed, T., Liu, M., Wu, Z., Luo, J., Tian, Y., et al., (2020). Identification of genes involved in antifungal activity of *Burkholderia seminalis* against *Rhizoctonia solani* using tn5 transposon mutation method. *Pathogens*, 9(10), 1–15. <https://doi.org/10.3390/pathogens9100797>.
- Zhao, L., Bao, Y., Wang, J., Liu, B., & An, L., (2009). Optimization and mechanism of diacetyl accumulation by *Enterobacter aerogenes* mutant UV-3. *World Journal of Microbiology and Biotechnology*, 25(1), 57–64. <https://doi.org/10.1007/s11274-008-9862-8>.
- Zhao, X. W., Zhou, M. H., Li, Q. B., Lu, Y. H., He, N., Sun, D. H., & Deng, X., (2005). Simultaneous mercury bioaccumulation and cell propagation by genetically engineered *Escherichia coli*. *Process Biochemistry*, 40(5), 1611–1616. <https://doi.org/10.1016/j.procbio.2004.06.014>.
- Zhou, Y., Wei, J., Shao, N., & Wei, D., (2013). Construction of a genetically engineered microorganism for phenanthrene biodegradation. *Journal of Basic Microbiology*, 53(2), 188–194. <https://doi.org/10.1002/jobm.201100322>.

CHAPTER 12

Molecular Markers and Their Applications

ABDUL REHMAN,^{1,2} HAFIZA IQRA ALMAS,³ ABDUL QAYYUM,⁴
HONGGE LI,² ZHEN PENG,² GUANGYONG QIN,¹ YINHUA JIA,²
ZHAOE PAN,² SHOUPU HE,^{1,2} and XIONGMING DU^{1,2}

*¹Zhengzhou Research Base, State Key Laboratory of Cotton Biology,
Zhengzhou University, Zhengzhou, China*

*²State Key Laboratory of Cotton Biology, Institute of Cotton Research,
Chinese Academy of Agricultural Science, Anyang, Henan, China*

³Department of Botany, University of Agriculture, Faisalabad, Pakistan

*⁴Department of Plant Breeding and Genetics, Bahauddin Zakariya
University, Multan, Pakistan*

ABSTRACT

Molecular markers perform a significant role in marker-assisted breeding (MAB) which is an advanced method of traditional breeding. It assists in the development of unique and new traits in plants within a short time. Moreover, it is used for tagging the distinctive features available at any site in the whole species genome. Many DNA and molecular markers are identified that regulate the change of genome sequence. Currently, QTL mapping increases the concepts relating to polygenes that regulate the vital characters of plants, whereas association mapping helps to provide a framework related to the phylogenetic traits. By the way, molecular

marker is significantly worked in this study. In this chapter, we sum up the molecular markers for enhancing the information related to their types and applications, in various features of breeding. Numerous markers are illustrated with their protocols and somewhat unique markers that are widely utilized today, such as CAPS, SNP, and dCAPS, etc. For agriculture research, extensive molecular genomic markers are present. These molecular markers are categorized into several groups based on their uses such as RAPD markers assist to find and screen the hybrids depending on drought and salt stress tolerance, although SSR markers are important for the determination of stress resistance. Similarly, these markers have a significant function in the QTL mapping of stress concerning genes. *Saltol* for salt stress is basic Gene that work under these stresses. Moreover, the SNP markers help in the mapping of genes and the sequencing of stress-associated characteristics in parent genotypes. DNA markers simplify marker-assisted breeding to increase the resistance in the response to abiotic stress by modern technologies and modifications of markers. In this section, we also describe the procedures used for different kinds of molecular markers that increase the possibility of their applications.

12.1 INTRODUCTION

Molecular breeding belongs to molecular biology; this method used for the breeding purpose (Moose & Mumm, 2008). The process in which the crossing over between the same species of the donor and recipient traits of plants is termed breeding, this approach is used for the betterment of quantitative and qualitative traits of the crop. Monogenic trait is defined as the desired traits developed by a single allele or gene, whereas polygenic is traits produced by multiple genes or alleles. With the support of this method, the breeding is extensively observed by trait locus mapping of whole genome. Different characters were studied by visual understanding; it is used in conventional breeding while modern methods are used for marker-assisted breeding (MAB). This process is analyzed through plants crossing over with the chosen character by cross-pollination. For diversification, the plants are interbred, for instance, the high protein content soybean hybrid is produced by crossing the plant species that have high content protein with the species that lack this quality. After and after, it

is backcrossed to produce high content protein traits. With the help of farmers, this process was performed decades ago, but today it comes to be a method for the development of potential resistance and better performance under adverse environmental conditions. Eventually, in developing countries, it becomes a critical approach to fulfill the needs of increasing population because the breeding method enhances the yield as well as the quality of crop.

12.2 CONVENTIONAL BREEDING

Roughly 1,000 years ago, farmers took part in altering the genetic structure of different crops. Before plant breeding, farmers selected the best seeds and plants by plant's visual appearance and preserved those seeds for the next germination season. Features including maximum production, resistance against diseases and pathogens, resistant against climate, and the size of seeds have been significantly modified from their wild families. For instance, on a large scale, the pattern of plant breeding has been observed successfully including the development of hybrid *Zea mays*, *Triticum aestivum*, and different genotypes of *Oryza sativa*, which produce green revolution (Duvick, 2001). Today, breeding has become simpler because of the initiation of these methods and the breeding techniques based on molecular markers (Agrawal, 1998). To study the traits and not only through phenetics consideration such as in conventional breeding, but MAB is also significant. MAB is build-up for analyzing the interaction of genotypes and phenotypes. Traits mapping by MAB has two groups: (i) QTL mapping; and (ii) association mapping. It consists of hybrid and mutation breeding technology of seeds.

12.3 MARKER-ASSISTED BREEDING (MAB)

Like traditional breeding, MAB includes the development and crossing of new genotypes through the transformation of specific features from a donor to a recipient plant with the help of molecular markers. Alleles of DNA markers related to the selected trait should be recognized for the confirmation of true plant. It does not contain the backcross ready for F8 progeny or the recovery of whole genome excluding the selected trait.

The molecular marker is utilized for the percentage recovery to analyze and acquires maximum improvement shortly. By MAB, whole-genome selection recovery is feasible to acquire in 3–4 backcross generations. Furthermore, it is very significant in QTL and association mapping. When plant representing number for the analysis of selected traits in specific population enhances exponentially, the number of QTL enhances but the efficiency of selection marker-assisted reduces (Moreau et al., 1998). As stated by Ribaut & Betrán (1999), mostly three QTLs are necessary for any MAB for significant results. Though, in many studies such as in tomatoes, five QTLs are also considered (Lecomte et al., 2004). Studies of QTL become feasible after the expansion of SNP marker selection now-a-day (Kumpatla et al., 2012).

12.4 QTL MAPPING

QTL is called quantitative trait locus. In whole-genome sequencing, it is a locus that describes the diversity in the quantitative trait of DNA sequence in species phenotype (Miles & Wayne, 2008). Traits that are controlled through multiple genes are known as the study of QTLs. For instance, several characters are mapped by QTL mapping in soybean (Kumawat et al., 2016). In the past, it was assumed that mapping of QTL includes the practice of both parents' populations mapping with minimum exposure of marker, however, it provided the poor resolution in the identification of QTL (Sonah et al., 2013). The study of QTLs initiates with the traits' identification that is authorized by the expression of the transient. But another is a candidate gene that depends on the homology of crops. Various QTLs have been identified and characterized thoroughly for essential genes and are also connected with haplotypes. Members of the leguminous family are comparatively dissimilar from other families of plants in several properties such as the production of many secondary metabolites, a high protein linked with the physiology and nitrogen fixation efficiency. Similarly, evolution exposes the genomes comparison and the methods that were not beneficial in the study of genetics (Koebner & Varshney, 2006; Varshney et al., 2015). Soybean is the model crop that belongs to a leguminous family that is utilized for the quantitative study of legumes due to the homology of all legume crops. In various leguminous crops excluding soybean *in silico* dependent study of homology that should precisely support the detection of the gene without the usage of QTL

mapping which is helpful in the sequencing of the gene. Advancement in molecular markers for specific features is useful for the genes detection in other legume crops (Kang et al., 2016). These are worked in the different types of population such as recombinant inbred genotypes, double haploid, backcross, and near-isogenic lines that all are resultant of both parents crossing for the required characters. Similarly, a huge inbreeds population and polymorphic markers are mandatory for mapping QTLs (Collard et al., 2005; Liu, 2017; Wu et al., 2007).

12.5 ASSOCIATION MAPPING

Usually, GWAS uses significantly efficient numbers of markers in a different genotype of population, which contain several events related to crossing-over from parents, support in high resolution than mapping of QTL and assistances in prediction and identifying the basic genes (Zhang et al., 2015). It is known as the study of phylogenetic trait that works in several organisms and their relationship. Phenotypic variation happened because of genetic variations studied by association mapping that gives defined linked alleles in common ancestors from all other species as well as helps in the identification of trait-linked polymorphism in several species. By the use of efficient markers such as SNPs, functional variations are detected (Hoeschele, 2004). It works in non-random loci in association mapping than the QTL mapping (Abdurakhmonov & Abdugarimov, 2008). It uses population depending on our analysis of the expected population or the source of germplasm, however, in QTLs crossing over between two parents are included. The following steps are present in association mapping. (i) An inclusive analysis of genetic variant experiment from the estimated population or the source of germplasm from natural species. (ii) The studies or determining the phenotypic properties such as crop tolerance and resistance against specific diseases, quality, and production of crop in various environmental circumstances and replications. (iii) Utilization of genotypic molecular marker for population mapping. (iv) Use of quantification for the analysis of the LD in the genome-wide population depends on the data of genotypic molecular markers. (v) To get the level of estimation of genetic variation by population structure between the experimental analysis and kinship among the individuals of the pairs controlled by any experiment. (vi) This step depends on the outcomes of the 5th step, such as the kinship, population structure, and kin between genotypic and

phenotypic information that exposes the position of associated marker proximity to the analysis of required trait by the ancestor for the study of phylogenetic accompanying by each other (Hansen et al., 2001; Kraakman et al., 2004, 2006; Nordborg et al., 2002). Association between gene and markers for disease resistance in various crops are provided in Table 12.1.

TABLE 12.1 Interaction between Markers and Genes in Various Crops for Disease Resistance

Species	Trait	Molecular Marker	Gene/QTLs/ Loci	References
Barley	Spot blotch resistance	DArT and SNP	Rcs5	Roy et al. (2010)
	Leaf rust	CAPS	Rph7	Graner et al. (2000)
	Fusarium head blight	SNP	vrs1	Massman et al. (2011)
	Barley yellow mosaic virus	SSR	rym4	Williams (2003)
Cotton	Verticillium wilt	SSR	CM12	Bolek et al. (2005)
	CIcuV	RFLP	pGH286b	Aslam et al. (2000)
Maize	Sugarcane mosaic virus	CAPS	Scmv1, 2	Dussle et al. (2002)
	Maize stalk rot	SSR	Rgsr8.1	Chen et al. (2017)
Rice	Gall midge	SCAR	Gm7	Sardesai et al. (2002)
	Rice blast	CAPS	Pi5	Jeon et al. (2003)
	Ergot resistance	DArT, SSR, and AFLP	SBI-02,06	Parh et al. (2008)
Sorghum	Anthraxnose resistance	SNP	Pib	Upadhyaya et al. (2013)
	Powdery mildew	SNP	QPm. stars-2BL1-3	Genqiao Li et al. (2019)
Wheat	Leaf rust	SSR	Lr34	Suenaga et al. (2003)
	Stem rust	STS	Sr31	Mago et al. (2002)
Sugar cane	Rust resistance	RFLP	CDSR29	Daugrois et al. (1996)
	Brown rust resistance	AFLP, gSSR, and EST-SSR	SCB100, 232	Santos et al. (2015)

12.6 MARKERS UTILIZED IN BREEDING PROCESS

In the genome, genetic diversity is responsible for the detection of traits that are linked to specific properties and polymorphism between two species. Characters transformed from one species to another are found by

the linked markers existing in that area. In plant breeding, the study of traits is transferred because it is the basic need of markers development. There are various types of markers such as biochemical, cytological, molecular, and morphological markers. On the base of base pairs, modification in the characters of two different sequences with a relative nucleotide in which specific DNA sequence site exposes molecular markers. Polymorphism develops because of various sequences in the similar site of the two characters in the analysis of base pair. In nucleotide sequences, the modifications exist through deletions, mutations, insertions, translocations, and duplications. Molecular markers exhibited benefits of distinctive polymorphisms that do not generate phenotypic diversity, extremely reproducible and can be useful for the genomic portion (Mondini et al., 2009). However, morphological markers can be selected phenotypically and qualitatively, such as height of plant, number of nodes, branches, leaf shape, flower color, and several other important field traits. But morphological markers have been not more helpful than the other markers because of the effect of environmental stresses. So, these markers are not effective (Eagles et al., 2001). Cytological markers that depend on the patterns of staining illustrate the variation in the position, size, shape, and order of chromosomes (Jiang, 2013). These differences provide the concept of the distribution of heterochromatin and euchromatin in chromosomes. Different patterns are formed through staining, such as G bands and Q bands are produced by Giemsa stain and quinacrine hydrochloride, respectively, in the study of the chromosome. This facilitates the interpretation of normal and mutated chromosomes. Biochemical markers have different biochemical proteins coded through gene that affects the functions of the trait. In the functions, this change is encoded through genes that are exposed by various markers to differentiate the characters from one another (Bailey, 1983). Similarly, these variations happen because of the change in the DNA sequence coding the gene through the mutation, such as frameshift, duplications, deletion, and inversion. There are some drawbacks of biochemical markers like effects of the environment, minimum quantity, and variation in different plant growth phases (Mondini et al., 2009).

12.7 DISTINCT FORMS OF MOLECULAR MARKERS

Different markers are recognized that depend on various methods and uses. Molecular markers such as RFLP, VNTR, SSR, dCAPS, CAPS,

RAPD, FLP, SNP, and AFLP. These markers are more beneficial than traditional breeding, because these depend on the selection of specific characters and visualization. Markers offer particular proof to help the trait associated with marker and consideration of different characteristics. These markers should have certain characteristics which become the good marker for evaluation such as polymorphic between two characters. According to research, several scientists supposed that RAPD marker has many advantages because of the minimum cost and reliable consequence, for instance, the study of RAPD markers on 269 common bean landraces indicates that the grouping result is same as in past which depends on ecology and morphology (Beebe et al., 2000). Furthermore, genetic variation in soybean genotypes is identified by RAPD that was initiated from China and America. Moreover, by RAPD markers, the maize population has been accurately determined (Li & Nelson, 2002; Popi et al., 2000). However, from agronomic and morphological data, the bands found in 42,000 sorghum germplasm accessions were associated with RAPD markers (Dahlberg et al., 2002). Hence, RAPD only one marker is not appropriate for the study of genetic variation. Similarly, species of wild rice were effectively categorized by RAPD markers (Farooq et al., 1995). These consequences show that RAPD marker is not reliable as well as uncertain. Hence, various molecular markers were utilized for breeding programs. AFLP markers were utilized for the analysis of 14 recognized strains of Bermuda grass, now cultivated in America (Anderson et al., 2001). In the same way, almost 105 cultivars of bread wheat were identified through AFLP and SSR markers. So, the SSR markers provide the difference in 105 varieties of bread wheat. But, the AFLP markers support in detection of wheat genetic variation in Argentina; similarly, both of these markers like SSR and AFLP are utilized for identifying the genetic variation in wheat genotypes that were produced in the past and current, but no reliable outcomes were obtained. However, through fingerprinting new cultivars of the narrow genetic base are produced that are easily detected through AFLP and SSR markers. The advantages of these markers exhibit that they are reliable and useful. For the analysis of germplasm, the RFLP is majorly utilized to give precise data on genetic variation (Andreas et al., 1994; Sasanuma et al., 1996; Siedler et al., 1994). It is mostly identified in many significant cases where sources of germplasm are used in the study of homology lineage. There are several reasons for their usefulness in the outcomes initiated by various markers that are used. One of the causes is

that the RFLPs are reliable markers to be utilized predominantly for the rule as illustrated in different crops such as species of wheat, sugarcane, rice, alfalfa, and some other crops (Oropeza & Degarcia, 1997; Pupilli et al., 2000; Sasanuma et al., 1996), however, they are not appropriate for the comparison of interspecific (Castagna et al., 1997). Genic molecular markers developed for various crops are given in Table 12.2.

12.7.1 RFLP

RFLP was the first recognized marker that depends on hybridization. By this marker, DNA polymorphism sequencing is detected in 1975. It is depending on the function of restriction digestion location that exists in sequences of DNA. Because of mutations such as insertion, duplication, deletion, DNA sequence modifies the identification spots existing in the original sequence. It produces the pattern of restriction digestion, while the mutated allele is recognized by this. RFLP determines the absence and presence of the mutation and also shows the condition of alleles such as heterozygous and homozygous because it is the co-dominant marker. It produces two separate digestion bands that show mutations and identify heterozygosity. By RFLP, the linked trait of several plant species can be easily recognized and also their maps have been made for other analysis (Tanksley et al., 1989). There are many advantages of RFLP such as maximum reproducibility, no requirement of PCR, co-dominant, high cost, abundance of high genomic, medium polymorphic, 10,000 ng DNA require, requirement of sequencing and visualization through radioactivity. RFLP is very useful to detect the larger sample size while RAPD is for sample of smaller sizes (Ragot & Hoisington, 1993). For full restriction, 2–10 µg DNA is required.

➤ Standardized Protocol:

- From the samples, the genomic DNA extracts;
- Restriction enzyme uses for digestion and restriction, this enzyme is isolated from the bacteria;
- On polyacrylamide gel electrophoresis (PAGE)/agarose, the larger bands separate;
- Shifting the separated bands on a nylon membrane/nitrocellulose;
- Labeled probes use for nucleic acid hybridization;
- Bands detect by autoradiography.

12.7.2 VNTR

Jeffreys et al. (1985) identified the VNTR markers. It exists in genome that has short tandem repeats in bands and are placed in similar directions. The bands of repeat indicate the difference in length among individuals that are depending on VNTR. These variations are accountable for the allele's detection. Also, it indicates the allele's existence or absence. It is observed through repeat block restriction digestion and in a gel. This marker is classified into two types such as micro and macro satellites. Microsatellite has <5 bp in length but microsatellite has longer blocked. Today, the SSR and STR are utilized which are very useful and helpful. Maximum 4 bp repeats are studied by VNTRs whereas, very short repeats such as 2 bp are failed that may change from one tissue to another in similar individuals. 3 bp repeats are analyzed to change from one progeny to another, for instance, in the disease of Huntington.

➤ **Standardized Protocol:**

- Restriction enzyme uses for the restriction digestion of repetitive sequence;
- Fragmented band analysis by gel electrophoresis.

12.7.3 SSR

Like other markers, this marker is also co-dominant. These markers depend on PCR. These are satellite markers that are present within the whole genome. They consist of di-nucleotide or trinucleotides repeat. They are highly reproducible, easy to identify, low DNA required and polymorphic. Whereas the identification cost is high because of the PCR-based in which several primers are used for whole-genome or specific parts of genome analysis. It is extensively utilized in every part of genetic and molecular breeding. SSRs have 1–6 nucleotides repeats. Hence, it may have mono, di, tri, and many more nucleotides repeat (Kalia et al., 2011). Their consistency is extraordinary, and they have minimum technical demands.

➤ **Standardized Protocol:**

- From desired samples, the genomic DNA separate;
- In PCR samples amplification by SSR primers having a thermal profile;
- PCR products separate on Metaphor agarose gel electrophoresis.

12.7.4 DCAPS

dCAPS is the co-dominant marker. It depends on restriction digestion and PCR. Both mutated and original samples are identified through the study of recessive and dominant alleles. Recessive alleles could be identified by amplified restriction products through the restriction enzyme. Mutant alleles could be detected by dCAPS markers containing restriction places for the restriction enzymes. These dominant alleles detect by single but larger base pair fragments, but recessive alleles containing genotypes will generate two smaller base pair fragments.

➤ **Standardized Protocol:**

- Separate the genomic DNA from the genotypes of parents and offspring;
- PCR multiplies the desired sample fragmented genes through the primers of dCAPS;
- Analysis of the PCR products by agarose gel;
- Purify and quantify the product of PCR;
- Two separate bands on the competition of restriction digestion (Watanabe et al., 2011).

12.7.5 CAPS

This marker is mainly utilized for the detection of recessive alleles. Amplification of desired gene fragmented PCR increases segment of DNA by almost 125 bp that have restriction area for Pdh1 enzyme. PCR Restriction digestion amplified the DNA that causes 76 and 49 bp fragments in genotypes that possess recessive Pdh1 enzyme. According to different researchers (Funatsuki et al., 2014; Spaniolas et al., 2006; Weiland & Yu, 2003), CAPS marker is co-dominant to classify the state of alleles such as homozygous or heterozygous and is extensively utilized in population genotyping. Similarly, it is worked in map-based cloning.

➤ **Standardized Protocol:**

- Separate and amplify DNA samples;
- Quantify the PCR product, clean the DNA and measure the amount of DNA;

- Prepare the reaction of restriction enzyme by the subsequent components of reaction;
- After the incubation, isolate the segments of the DNA on 3% agarose gel. Take the picture of gel. Genotypes by Pdh1 will generate the single 125 bp segments fragment, while genotypes having recessive gene pdh1 will generate the 76 and 49 bp segments.

12.7.6 RAPD

RAPD is mostly a PCR-based marker. It is dominant marker that do not differentiate the state of the allele. Similarly, it does not need previous data for the sequence of the genome. Several random primers are needed (approximately size 8–12 nucleotides) for binding arbitrarily in the genomic sequence of offspring and parent genotypes. Binding of primer is based on the possibility of corresponding nucleotide of genome. RAPDs were capable of consecutively identifying various loci of polymorphisms within the whole genome in a single cycle (Williams et al., 1990). High reproducibility and polymorphic, dominant, medium amount of DNA required, high abundance of genomic, low-cost, data of sequencing not needed only arbitrarily primer attaches and amplification happens, requirement of PCR, agarose for visualization, DNA only 20 ng required. This marker is not reliable but requires low cost and technical demands.

➤ Standardized Protocol:

- Segregate the genomic DNA from parents genotypes and offspring;
- Do amplification with PCR by the short arbitrary primers including nucleotides 8–12 bp;
- Analysis of the PCR product in gel electrophoresis.

12.7.7 AFLP

It includes different methods such as PCR and restriction and gives reproducible outcomes. Medium reproducibility, dominant, maximum polymorphic, high amount of DNA needed, such as 500 to 1,000 ng. PCR-based, high cost, and genomic abundance, no sequencing required, agarose gel required for visualization. Zabeau (1993) illustrated that

AFLP is a PCR-based technique utilized in the fingerprinting of DNA. It is a complex process and useful in polymorphism detection within the sequence of DNA. About 500 ng DNA is needed, while it is a very helpful marker with intermediate cost and technical needs.

➤ **Standardized Protocol:**

- Restriction enzyme is used for the digestion of genomic DNA;
- After, the ligation of adaptor sequence within the half-site restriction;
- Two cycles of amplification of PCR are completed with the extension of primer such as 1 bp and 2 bp, respectively, which have adaptors and particular sites for restriction sequences;
- PCR products are visualized by gel electrophoresis.

12.7.8 FLP

This marker has been acquired from these great deletion/insertion sites that can be utilized for the method of molecular breeding. Several types of research have been performed by Watanabe et al. (2011); Liu et al. (2008); and Xu et al. (2013).

➤ **Standardized Protocol:**

- From desired samples, the genomic DNA separates;
- The desired gene segments amplified by PCR with the help of FLP primers and thermal profile;
- Analysis of the PCR product in agarose gel electrophoresis;
- Genotypes in the dominant form will make the single 1,339 bp fragment such as in soybean the E3 gene, while genotypes in recessive form will generate different sizes of fragments, such as smaller 558 and larger 275 bp.

12.7.9 SNP

It is a sequence-based marker in which single base pair alter in the same site in two individuals that make the SNP. Either it modifies or not the identification places for restriction enzyme hence, it includes the site-sequencing to differentiate between two individuals. Designing of primer

depends on the single base pair variation. A single base pair variation in the genome sequence produces the SNP. The SNPs may be changed the purine known as transitions, i.e., T/C and A/G while pyrimidine-purine variations called transversions, i.e., G/C, EST, and C/A. These markers contain a single bp variation, such as deletions/insertion modifies the coded sequence of mRNA. It is a highly plentiful marker available in the whole genome because of the single bp modification. Different ranges of SNP in plants, such as 100 to 300 base pairs (Xu, 2010). In advancement, the SNPs markers are getting various applications in molecular markers. Highly polymorphic, co-dominant, and reproducibility, almost 50 ng DNA required, genomic highly abundant, different cost, required PCR. The SNP-VISTA is used for visualization (Mammadov et al., 2012; Rafalski, 2002). According to Bhatramakki et al. (2002); and Jones et al. (2007), the SNP markers provide the mapping of high-resolution than other molecular markers. Wright et al. (2005) researched the SNP markers through re-sequencing Uni-Genes resultant of the amplicons by Sanger's method. Batley et al. (2003) studied the *in-silico* detection of the SNP by the SNP-mining. Databases of EST are developed by the verification of PCR.

TABLE 12.2 Genic Molecular Markers Developed for Various Crops

Plant Name	Species	Marker Type	References
Finger millet	<i>Eleusine oracana</i> L. Gaertn	EST-SSR	Nirgude et al. (2014)
Chickpea	–	EST-SSRs, EST-ITPs, EST-ESTPs, EST-SNPs	Choudhary et al. (2012)
Pea nut	<i>Arachis hypogea</i> L.	EST-SSRs	Mondal et al. (2012); Liang et al. (2009)
Linseed	<i>Linum usitatissimum</i> L.	EST-SSRs	Soto-Cerda et al. (2011)
Sweet potato	<i>Ipomoea batatas</i>	EST-SSRs	Wang et al. (2011)
Pea	<i>Pisum sativum</i> L.	EST-SSRs, EST-CAPS	Jain et al. (2014)
Castor bean	<i>Ricinus communis</i> L.	EST-SSRs	Qiu et al. (2010)
Sorghum	<i>Sorghum bicolor</i> L.	ILP	Jaikishan et al. (2015)
Cotton	<i>Gossypium</i> spp.	EST-SSRs	Park et al. (2005)
Wheat	<i>Triticum aestivum</i> L.	EST-SSRs	Nicot et al. (2004)
Maize	<i>Zea mays</i> L.	cDNA-RFLP	Falque et al. (2005)
Safflower	<i>Carthamus tinctorius</i> L.	IFLP, EST-SSR, RGC	García-Moreno et al. (2010)
Sunflower	<i>Helianthus annus</i> L.	EST-SNP	Lai et al. (2005)

12.8 MOLECULAR MARKERS APPLICATIONS

Several molecular markers are useful for genetic studies of any character in a different generation, for the analysis of phylogenetics in any organism. Furthermore, it is beneficial for the identification of relationships among different and same species. These markers are helpful for the detection of hereditary diseases, genotyping, forensic sciences, and paternity testing. Similarly, in the agronomic field, the genes identified that are responsible for tolerance and resistance to disorder can be detected. It is also valuable for the study of recombinant frequency and genetic mapping in all progenies. Moreover, these markers are appropriate for the detection of micro-organisms. Simultaneously, AFLP help to detect the various polymorphisms in the genome of different species where genetic composition is unidentified of bacteria, plants, and fungi. This tool is very useful in the study of genetic diversity taxa in closely related species of bacteria, fungi, and plants. Also, it is a complex and highly reproducible method for the study of linkage and production of QTLs map. For the linkage map assembly, these markers are very beneficial and suitable techniques. These are useful for the study of handling of disorders, somaclonal variation, and detection of genetic distance in the specific traits. Commonly, it is the basic technique used for different methods of molecular breeding (Table 12.3).

12.9 APPLICATIONS OF RAPD MARKERS UNDER ABIOTIC STRESS TOLERANCE

RAPD markers are PCR-based. Previous sample sequencing information is not mandatory for the study of RAPD. Several loci from various individuals can be studied for selection targets through inadequate sources. These markers are extensively utilized because of their easy practical procedures and tremendous genetic screening of hybrids such as inter-specific and intraspecific (Shasany et al., 2005). RAPDs have helped to find the genes associated with tolerance against salinity stress in various plants. Many defensive mechanisms are present in plants under salt stress. These pathways are controlled genetically. Hence, developing tolerance against salt stress in many crops is complex and sensitive, particularly in saline-affected regions. DNA markers can support finding and classifying salinity tolerant varieties. Using PCR for RAPD marker amplification of particular DNA sequences is an important technique for identifying the

salinity-tolerant genes. Research was performed in wheat to identify the genetic variation of salt-tolerant genotypes utilized by plants producing within saline-affected areas. RAPD markers efficiently discriminated the salt-tolerant from salt-sensitive varieties. The pairing of Polymorphic primer between resistant and sensitive varieties approved the genetic diversity. Resistant wheat genotypes might be produced by such genetic data to cross between the salt-tolerant and salt-sensitive varieties (Bhutta & Amjad, 2015). Several modifications in DNA can be produced under salt stress, including rearrangement and breakdown of structure. These modifications are because of secondary stress, such as oxidative destruction associated with the development of ROS, including singlet oxygen, superoxide, hydrogen peroxide, and hydroxyl radicals. These markers are useful to detect the genetic variability in cotton plants growing under saline-affected fields like the treatment of NaCl. These markers exposed the missing bands, strength of strong and weak bands, and existence of new bands than control plants of DNA in agarose gel electrophoresis. Prior research proved that RAPD uses efficiently detected the stress related to toxicology. Primer of RAPD was useful and had a considerable capacity for detecting the changes in DNA variations by salinity stress. Awkwardly, many challenges still present for the uses of RAPD. These challenges consist of amplification features and separation of electrophoresis including DNA impurity, the existence of identical bands and competition within the amplification of DNA (Doğan et al., 2012). Field estimation is relatively difficult and time taking for categorizing the factors of quality, stress-tolerant, and yield of crop. At the DNA stages, molecular transformations were effectively classified in plants cultivated from plant tissue culturing (Jain, 2001). In stress conditions, plant genetic screening has been done through molecular markers technique (Balkrishna & Shankarrao, 2013). RAPD is simple and fast and needs a small quantity of DNA. This tool is critical to genetic modifications and can rapidly process huge genomic samples within *in vitro* situations. For instance, in maize polymorphic bands indicating salt-tolerant genes were studied through RAPDs. Many primers were associated with salt-tolerant sequences which delivered useful data for the breeding of salt-tolerant progeny. These generations can be more screened via a selection of marker-assisted and produced for salt tolerance through direct genetic differences of varieties (Karp, 1997). GTS (genomic template stability) evaluation is a rate of DNA damage and DNA mutation such as point mutation, frame-shift mutation, and

structural mutation. Bands of the DNA obtained by the study of RAPD can successfully participate in the construction of the molecular marker for the detection of mutant and damaged DNA in plant cells. GTS is identifying the differences qualitative in RAPD profiles. Research on cotton varieties including A118, DE22, N78, and DP50 was performed to calculate the genotypes response grown in different conditions such as under 200 mM NaCl stress and non-saline. Modifications in the RAPD profile were calculated through GTS in the form of percentage. A significant result was obtained in two sensitive varieties such as A11f (51.2%) and DP50 (79.1%). On the other hand, the least values were found in two salt-tolerant genotypes including DE22 (36.7%) and N78 (26.4%). Similarly, DNA modification might be used for developing resistant germplasm against salinity stress in plant breeding programs. Furthermore, salt-tolerant varieties such as DE22 and NB78 holding the minimum GTS values indicated the maximum RAPDs polymorphic expression. According to these consequences, the study of RAPD helps to detect the sequences of DNA associated with salinity stress. Therefore, these molecular markers might give the discoveries in the earliest identification of salt-tolerant varieties for the cotton crop (Saleh, 2016). Due to drought, the yield and quality of wheat are significantly influenced in various areas. Losses in drought are the same as losses because of different other environmental stresses. This situation is worsened by global environmental stresses (Kirigwi et al., 2007; Shao et al., 2005). Expressions of drought-induced in several genes participate directly in stress tolerance (Zhou et al., 2010). Drought-tolerant-associated primers of the DNA were used in preliminary DNA fingerprinting for the study of RAPD to observe the genetic diversity in wheat genotypes. A similar analysis required the genetic details for drought-tolerant in hybrids of wheat. A P6 RAPD primer (TCGGCGGTTC) amplified the 920 bp band under semi-drought that was deficient in drought-tolerant genotypes. Moreover, a 717-bps band of RAPD such as the *Dreb-B1* gene was found in the genome of B resultant of drought-tolerant like Barakatli-95 (Huseynova & Rustamova, 2010). Rashed et al. (2010) observed 4 positives and 2 negatives in RAPD markers which proved the reliability of RAPD for the detection of drought-tolerant varieties of wheat. In gardening plants such as tomato plants are produced less yield in high temperatures. Features related to productivity are estimated to be inherited quantitatively and mainly influenced by the alterations within climatic stress situations. The complexity features develop difficulty to measure the crops, particularly

in heat tolerance. RAPD analysis associated with heat-tolerant genotypes of tomatoes in high-temperature conditions recognized 43 recombinant hybrids lines in F7 progenies produced by a crossing of wild and heat-tolerant parents including L4422 and CL 5915. Only 14 among 200 RAPDs markers were recognized as indicating heat-resistant by bulked analysis of segregant. Some RAPD were particular for only one trait, but many were linked with two. RAPD associated with heat tolerance indicated the significant effect of the gene because of the CL5915 gene. Molecular markers of genetic screening particularly for heat-resistance can be improved by selecting an individual variety with targeted characters marker (including weight, productivity, and number associated markers) like in tomato crop D11, P06, D06, and X01 markers that may provide conventional breeding by CL5915 as a donor parent (Lin et al., 2006).

12.10 HYBRIDS AND STRESS TOLERANCE

Screening of marker-assisted is a useful method for enhancing abiotic tolerance in crops. Detection of stress-tolerant genes is usually based on SSR markers. Analysis of SSR with bulk segregant is efficient for detecting the DNA markers associated with agricultural traits including heat-resistance and grain filling in wheat plants. Three markers namely Xgwm617, Xgwm577, and Xgwm132 were recognized through the analysis of SSR. These markers are related to the grain filling amount in high-temperature conditions. These methods assisted in the genotypes' development with enhanced heat-tolerant stress. Also, in rice, SSR markers including RM3586 at chromosome 3 while RM3735 at chromosome 4 showed a positive association with heat-tolerance expressing 3% and 7% total genetic diversity, respectively (Ashraf & Harris, 2005; Najeb et al., 2011; Zhang et al., 2009). These markers helped in the detection of drought-tolerant tetraploid hybrids of cotton. Allelic polymorphic consequences by SSR primers and agronomic traits expressed important results in Sayar-314 and Varamin hybrids, while hybrids of Tabladila exposed significant polymorphic information through the markers of EST-SSR. Hybrids of drought-resistant, like Nazily indicated 53% polymorphism (Poortavakoli et al., 2017). Detection of genetic variation in frost and heat stress selection for heat-tolerant cultivars in field environments, such as morphological selection is not better because of uncontrollable environmental effects which compromise the precision and repeatability of trial. Besides, the

definite predictability of heat-stress in field zones is not feasible. Genetic evaluation of quantitative properties for adaptive responses is required. The molecular study approves the uses of particular cultivars in breeding approaches for improving the stability of productivity and sustainability of crops in stressful conditions (Collins et al., 2008).

12.10.1 GENETIC DIVERSITY IDENTIFICATION UNDER ABIOTIC STRESS

Heat tolerance is having multi-genetic features with several mechanisms of resistance controlled by different gene clusters in various tissues and growth phases. SRAP markers are PCR-based that recover the segments of DNA in only one reaction of PCR. These molecular markers amplify several reproducible and polymorphic loci and alleles. SRAP markers can amplify particular, active, and efficient genes since the sequence of genes are mandatory for different processes. Based on multi-locus and multi-allelic nature, SRAPs are sometimes used for the fingerprinting of the DNA, mapping of genes and estimation of genetic variation. Arbitrary distribution in the genome of a plant is not appropriate for the application of SRAPs (Genyi & Quiros, 2001). Similarly, TRAP markers are effective and active PCR-based that works with two 18-mer DNA primers. In which one primer is fixed with EST while the other is associated with GC/AT abundant core to link with an intron/exon (Bohnert et al., 2006; Howarth, 2005). TRAP markers were used on wheat cultivars cultivated under heat tolerance. Genetic studies were conducted for TRAP and SRAP markers to estimate the genetic variation in durum wheat varieties. Genetic variation in agricultural traits against heat stress was observed. Information retrieved from field depends on the agricultural characteristics utilized in complex and multi-genetic forms. The methods of molecular markers depend on the selection of generation for the stress-tolerant traits. On the other hand, frost is badly affected crop productivity and death in pea plants. The agricultural studies to enhance the frost-tolerance in peas utilized 672 various genotypes of pea from three various regions. A trait-dependent marker link was utilized to evaluate the frost-tolerance with 267 SSR markers. Of these cultivars, only 16 were identified as frost-resistant that had the potential to survive in all field trials in that research. Structure of population expressed a population structured of two sub-populations with some hybrids in the 672 cultivars. An association approach detected 7 SSRs markers which

frequently expressed the association with winter-tolerance in the least two various climatic conditions by two statistical patterns. EST-1109 is one marker on LG VI which was expected to localize through a gene that takes part in the metabolism of glycoproteins under frost stress conditions. This gene prompts various mechanisms for frost tolerance in pea plants. These frost-tolerance genotypes and winter-tolerance associated markers act a major function in breeding through marker-assisted for winter-tolerant genotypes of peas (Liu et al., 2017).

12.11 QTL GENES MAPPING BY DNA MARKER SALTOL

Rice chromosomes have been mapped and salt-tolerant QTLs have been discovered. Saltol, which is found on rice chromosome 1, is one of the most important QTLs. Following saltol cloning site, the gene was detected which was associated with minimum absorption of sodium ions and high uptake of potassium ions. This gene outcome in a low ratio of Na and K phenotype in high salt stress (De Leon et al., 2015). For gene mapping, SSR especially microsatellites are appropriate since such molecular markers have been used to assess the genetic variation in several crops. Their simple method, co-dominance, reproducibility, and huge potential to distinguish the polymorphisms are chosen to estimate the profiles of DNA of plant species. Estimation of progeny, phylogenetic study, genome mapping and genetic differences is possible by these DNA markers. Research showed the genetic variation among rice cultivars by SSR markers. These fragments of DNA are a significant method for measuring the salinity tolerance in rice germplasms. The results of the research addressed differences among SSRs associated with salinity-tolerant plants such as Pokkali, a significant expressional cause of saltol. On chromosome no. 1, saltol QTL was mapped by using cross hybrids of 8 generations (Pokkali $\times$ IR29). Among 33 SSRs, the primer of RM8094 was discovered to be important in genetic variation. Analysis of cluster classified the cultivars into 3 groups that were included 16, 8, and 12 cultivars. Maximum salt tolerant IRRI screened varieties were assembled to classify into one group. Additionally, the salt-tolerant and intermediate-tolerant lines such as Pokkali and FL478 varieties were grouped into 2nd class. Highly resistant and sensitive lines were distributed individually into 3rd class. RM10745 and RM8094 expression of molecular markers were suitable for classifying the hybrids of salinity-tolerant (Mohammadi-Nejad et al., 2008). The gene of saltol was

transferred into a widely held rice genotype through breeding of marker-assisted by the selection of phenotypic and genotypic. The SSRs namely RM493/RM3412b were valuable for the selection of gene saltol. Gene of saltol introgression into recipient genotypes can be done by co-dominant SSRs. FL478 was main parent contributor for the effective salt-tolerant transmission into the genome progeny of BT7 (Linh et al., 2012).

12.12 QTL MAPPING FOR ABIOTIC STRESS BY DNA MARKERS

The main abiotic stress is drought which can reduce the yield in various wheat-producing regions globally. Molecular markers related to QTLs, particularly for drought tolerance might be significantly enhanced the drought-resistant in hybrids. Research on the QTLs identification related to genes that produced grain in drought stress conditions and give significant data about genes. STSs were utilized for mapping the QTL for hybrid lines of wheat drought tolerant. Grain productivity QTL is observed in the proximal chromosomal site 4AL. This site is related to the quantity of grain filling, spikes density, productivity of grain and biomass, and drought sensitivity index (Kirigwi et al., 2007). Genetic variation is the highest in rice because of ancestor species and extensive plant distribution in several areas. Several stresses decrease the productivity of the crop. In yield, almost 50% of losses are because of abiotic stresses only. Salt stress is an important biophysical issue for the development of rice in several areas (Munns & Tester, 2008). Salt stress tolerance is complex physiological and genetically mechanism. Different QTL associated with stress was detected in rice. QTLs identification for salinity tolerance with densely related proximate molecular markers might be suitable for coping with traditional breeding approaches that are deeply based on morphological analysis (Sabouri & Sabouri, 2008). In rice varieties, 20 QTLs were identified on chromosomes 01, 02, 04, 06, 08, 09, and 12. Unique QTLs like qSESI12.1 and qSESF12.1 might be used to narrow down locus mapping and discover closely related flanking markers for improving salt tolerance (Bizimana et al., 2017). The yield and stability of pea are severely affected by drought stress. Fewer researchers reported on drought resistance and linked genetic sources of *Pisum sativum*. One research was conducted on genomic loci associated with drought resistance. Symptoms of drought and leaves and soil relative water content under drought stress were evaluated for recombinant pea hybrid lines produced from two

parents identified to screen for drought-tolerance 10 QTLs related to characteristics independently relating from 9–33% variation of morphology. Reproducible DNA markers were detected to link with QTLs. These DNA markers may be used to select the individuals that show the desired QTLs in breeding *Pisum sativum* for drought resistance (Iglesias-García et al., 2015). Expanding cold resistance for the development of cold season genotypes of peas is an important issue. Inbreed of cold season genotypes require to consider cold tolerance along with quality and productivity of seed. Genetic factors of freezing tolerance and detection of genetic relation with production and yield characteristics were the objective of one research. A unique identified basis of cold resistance was utilized, and recombinant hybrid lines population was estimated in six various environmental stresses. The genetic map containing 679 DNA markers around seven linkage clusters and including 947.1 cM was produced. Around 161 QTLs responsible for 9–71% of morphological variations were detected for all considered characters. Results expressed that cold tolerance may be produced independently to increase the quality and production of seed (Klein et al., 2014).

TABLE 12.3 Utilization of DNA Markers in Field Crops for QTL Mapping

Crop	DNA Marker	Abiotic Stress	Number of QTLs	Chromosome	References
Barley	SSR	Salt	13	4	Xue et al. (2009)
	SNP	Salt	02	1 and 2	Sbei et al. (2014)
Cotton	SSCP	Drought	14	11	Abdelraheem et al. (2018)
	SSR	Salt	—	—	Saeed et al. (2014)
Maize	SNP	Salt	29	1, 2, and 5	Wan et al. (2017)
Rice	SSR	Drought	47	4 and 12	Babu et al. (2003)
	SSR	Heat	02	4 and 10	Xiao et al. (2011)
	SSR	Heat	01	3	Buu et al. (2014)
Sorghum	RFLP	Drought	07	1 and 2	Kebede et al. (2001)
Wheat	SSR	Drought	01	2	Dolferus et al. (2019)
	SNP	Salt	06	7	Hussain et al. (2017)
	SNP	Heat	03	3, 4, and 5	Tadesse et al. (2019)
	SSR	Cold	02	8 and 10	Wainaina et al. (2018)

Note: RFLP: Restriction fragment length polymorphism; SNP: Single nucleotide polymorphism; SSR: Single sequence repeat; SSCP: Single strand conformation polymorphic.

12.13 SNP MARKER AS A MARKER-ASSISTED SELECTION

In a huge population, association mapping evaluates more alleles as compared to linkage analysis. The mapping indicates the advantage of detecting phylogenetic recombination and several different lines having mutational features. Particularly, this process detects the genes related to phenotypic variation. Recently, the study of gene mapping applications is more significant for gene recognition accounting for the quantitative differences complicated traits such as drought tolerance (Yan et al., 2011; Zhu et al., 2008). In contrast, in plant species association mapping is inadequate for detecting the exceptional alleles. As well as, association mapping has a high cost because of the high genotyping and sequencing lines required (Yan et al., 2011). The use of fixed complex SNP chips is low cost and time-saving for linkage of genome-wide and association mapping. Conversely, analysis of linkage needs segregation of alleles through multi-allelic markers and capillary electrophoresis (CE). Earlier researchers prove that chips of SNP provide comprehensive genetic data, quality information, and accuracy in genotyping data. Furthermore, the SNP markers may be better for the analysis of linkage than conventional molecular markers, i.e., SSR. The SNP markers are advanced DNA markers, available in abundance relating to diversity. Genetic variation and functional genes are detected through SNP because of wide-genome assembly features (Yan et al., 2010). Hao et al. (2011) found 27 SNPs associated with drought tolerance genetic diversity in maize lines by identifying functional gene variations.

12.14 FUTURE OUTLOOKS

Our knowledge respecting molecular markers simplified the multiple characters related to linked markers. The markers analysis is significant in the breeding method because the molecular marker is reliable for the specific traits that are ultimately linked with importance while these are beneficial for the growth of the plant. We have concise the methods of breeding and various kinds of markers utilized in breeding that helps to improve the growth of plant populations. Through molecular markers, several QTLs are recognized. Similarly, wide genomic sources and molecular markers are utilized for the detection of important characters. MAB provides the chance to examine and get modifications in the earlier unconcerned regions of plant sciences and further gets the genetic sources for extensive

plant-growing methods. These processes are significant for the proper uses of crops and various plant species in a valuable approach in the future. In crops, the handling of abiotic stress is required for better quality and high production. Molecular genetics has many molecular markers that identify genetic variation, genotypic tolerance, stress-resistant lines, and genetic data associated with abiotic stresses. In the beginning, the molecular marker tools offered DNA markers which provided the basic data for stress tolerance. Conversely, present-day uses modern markers can detect the particular genes or genome reliable for abiotic stress resistance. Both mappings of QTLs and DNA show the arrangement of stress resistance genes on particular loci of the chromosome. Hence, in the tool of molecular markers, the need for continuous enhancement will permit even more comprehensive studies of stress resistance in climatic variations. Currently, in various world laboratories, research on these technologies is proceeding. In the future, tools of sequencing with minimum cost will support the detection of advanced genes in minimum time in several crop species with successful and efficient methods.

KEYWORDS

- **genomic template stability**
- **marker-assisted breeding**
- **markers**
- **plant breeding**
- **polymorphism**
- **quantitative trait locus**
- ***Zea mays***

REFERENCES

- Abdelraheem, A., Fang, D. D., & Zhang, J., (2018). Quantitative trait locus mapping of drought and salt tolerance in an introgressed recombinant inbred line population of upland cotton under the greenhouse and field conditions. *Euphytica.*, 214(1), 1–20.
- Abdurakhmonov, I. Y., & Abdukarimov, A., (2008). Application of association mapping to understanding the genetic diversity of plant germplasm resources. *International Journal of Plant Genomics*, 2008.

- Agrawal, R. L., (1998). *Fundamentals of Plant Breeding and Hybrid Seed Production*. Science Publishers, Inc.
- Anderson, M., Taliaferro, C., Martin, D., & Anderson, C., (2001). Comparative DNA profiling of U-3 turf bermudagrass strains. *Crop Science*, 41(4), 1184–1189.
- Ashraf, M., & Harris, P., (2005). *Abiotic Stresses: Plant Resistance Through Breeding and Molecular Approaches*. CRC Press.
- Aslam, M., Jiang, C., Wright, R., & Paterson, A., (2000). *Identification of Molecular Markers Linked to Leaf Curl Virus Disease Resistance in Cotton*. 11(4), 277–280.
- Babu, R. C., Nguyen, B. D., Chamarek, V., Shanmugasundaram, P., Chezian, P., Jeyaprakash, P., Ganesh, S., et al., (2003). Genetic analysis of drought resistance in rice by molecular markers: Association between secondary traits and field performance. *Crop Science*, 43(4), 1457–1469.
- Bailey, D. C., (1983). Isozymic variation and plant breeders' rights. In: *Developments in Plant Genetics and Breeding* (Vol. 1, pp. 425–440). Elsevier.
- Balkrishna, R. A., & Shankarrao, S. S., (2013). *In vitro* screening and molecular genetic markers associated with salt tolerance in maize. *African Journal of Biotechnology*, 12(27), 4251–4255.
- Batley, J., Barker, G., O'Sullivan, H., Edwards, K. J., & Edwards, D., (2003). Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. *Plant Physiology*, 132(1), 84–91.
- Beebe, S., Skroch, P. W., Tohme, J., Duque, M., Pedraza, F., & Nienhuis, J., (2000). Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. *Crop Science*, 40(1), 264–273.
- Bhatramakki, D., Dolan, M., Hanafey, M., Wineland, R., Vaske, D., Register, J. C., Tingey, S. V., & Rafalski, A., (2002). Insertion-deletion polymorphisms in 3' regions of maize genes occur frequently and can be used as highly informative genetic markers. *Plant Molecular Biology*, 48(5), 539–547.
- Bhutta, W. M., & Amjad, M., (2015). Molecular characterization of salinity tolerance in wheat (*Triticum aestivum* L.). *Archives of Agronomy and Soil Science*, 61(11), 1641–1648.
- Bizimana, J. B., Luzi-Kihupi, A., Murori, R. W., & Singh, R., (2017). Identification of quantitative trait loci for salinity tolerance in rice (*Oryza sativa* L.) using IR29/hasawi mapping population. *Journal of Genetics*, 96(4), 571–582.
- Bohnert, H. J., Gong, Q., Li, P., & Ma, S., (2006). Unraveling abiotic stress tolerance mechanisms—getting genomics going. *Current Opinion in Plant Biology*, 9(2), 180–188.
- Bolek, Y., El-Zik, K. M., Pepper, A. E., Bell, A. A., Magill, C. W., Thaxton, P. M., & Reddy, O. U. K., (2005). Mapping of verticillium wilt resistance genes in cotton. *Plant Science*, 168(6), 1581–1590. doi: <https://doi.org/10.1016/j.plantsci.2005.02.008>.
- Buu, B. C., Ha, P. T. T., Tam, B. P., Nhien, T. T., Van, H. N., Phuoc, N. T., Giang, L. H., & Lang, N. T., (2014). Quantitative trait loci associated with heat tolerance in rice (*Oryza sativa* L.). *Plant Breeding and Biotechnology*, 2(1), 14–24.
- Castagna, R., Gnocchi, S., Perenzin, M., & Heun, M., (1997). Genetic variability of the wild diploid wheat *Triticum Urartu* revealed by RFLP and RAPD markers. *Theoretical and Applied Genetics*, 94(3, 4), 424–430.
- Chen, Q., Song, J., Du, W. P., Xu, L. Y., Jiang, Y., Zhang, J., Xiang, X. L., & Yu, G. R., (2017). Identification, mapping, and molecular marker development for Rgsr8. 1: A new

- quantitative trait locus conferring resistance to *Gibberella* stalk rot in maize (*Zea mays* L.). *Frontiers in Plant Science*, 8, 1355.
- Choudhary, S., Gaur, R., Gupta, S., & Bhatia, S., (2012). EST-derived genic molecular markers: Development and utilization for generating an advanced transcript map of chickpea. *Theoretical and Applied Genetics*, 124(8), 1449–1462. doi:10.1007/s00122-012-1800-3.
- Collard, B. C., Jahufer, M., Brouwer, J., & Pang, E., (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, 142(1), 169–196.
- Collins, N. C., Tardieu, F., & Tuberosa, R., (2008). Quantitative trait loci and crop performance under abiotic stress: Where do we stand? *Plant Physiology*, 147(2), 469–486.
- Dahlberg, J., Zhang, X., Hart, G., & Mullet, J., (2002). Comparative assessment of variation among sorghum germplasm accessions using seed morphology and RAPD measurements. *Crop Science*, 42(1), 291–296.
- Daugrois, J. H., Grivet, L., Roques, D., Hoarau, J. Y., Lombard, H., Glaszmann, J. C., & D'Hont, A., (1996). A putative major gene for rust resistance linked with a RFLP marker in sugarcane cultivar 'R570'. *Theoretical and Applied Genetics*, 92(8), 1059–1064. doi: 10.1007/BF00224049.
- De Leon, T. B., Linscombe, S., Gregorio, G., & Subudhi, P. K., (2015). Genetic variation in Southern USA rice genotypes for seedling salinity tolerance. *Frontiers in Plant Science*, 6, 374.
- Doğan, İ., Kekeç, G., Özyiğit, İ. İ., & Sakçalı, M. S., (2012). *Salinity Induced Changes in Cotton (Gossypium hirsutum L.)*. *Pakistan Journal of Botany*, 44(SPL.ISS.1), 21–25.
- Dolferus, R., Thavamanikumar, S., Sangma, H., Kleven, S., Wallace, X., Forrest, K., Rebetzke, G., et al., (2019). Determining the genetic architecture of reproductive stage drought tolerance in wheat using a correlated trait and correlated marker effect model. *G3: Genes, Genomes, Genetics*, 9(2), 473–489.
- Dussle, C., Quint, M., Xu, M., Melchinger, A., & Lübberstedt, T., (2002). Conversion of AFLP fragments tightly linked to SCMV resistance genes Scmv1 and Scmv2 into simple PCR-based markers. *Theoretical and Applied Genetics*, 105(8), 1190–1195.
- Duvick, D. N., (2001). Biotechnology in the 1930s: The development of hybrid maize. *Nature Reviews Genetics*, 2(1), 69–74.
- Eagles, H. A., Bariana, H. S., Ogonnaya, F. C., Rebetzke, G. J., Hollamby, G., Henry, R. J., Henschke, P., & Carter, M., (2001). Implementation of markers in Australian wheat breeding. *Australian Journal of Agricultural Research*, 52(12), 1349–1356.
- Falque, M., Décousset, L., Dervins, D., Jacob, A. M., Joets, J., Martinant, J. P., Raffoux, X., et al., (2005). Linkage mapping of 1454 new maize candidate gene loci. *Genetics*, 170(4), 1957–1966.
- Funatsuki, H., Suzuki, M., Hirose, A., Inaba, H., Yamada, T., Hajika, M., Komatsu, K., et al., (2014). Molecular basis of a shattering resistance boosting global dissemination of soybean. *Proceedings of the National Academy of Sciences*, 111(50), 17797–17802.
- García-Moreno, M. J., Velasco, L., & Pérez-Vich, B., (2010). Transferability of non-genic microsatellite and gene-based sunflower markers to safflower. *Euphytica*, 175(2), 145–150. doi: 10.1007/s10681-010-0139-6.

- Graner, A., Ludwig, W. F., & Melchinger, A. E., (1994). Relationships among European barley germplasm: II. Comparison of RFLP and pedigree data. *Crop Science*, 34(5), 1199–1205.
- Graner, A., Streng, S., Drescher, A., Jin, Y., Borovkova, I., & Steffenson, B., (2000). Molecular mapping of the leaf rust resistance gene Rph7 in barley. *Plant Breeding*, 119(5), 389–392.
- Hansen, M., Kraft, T., Ganestam, S., SAeLL, T., & Nilsson, N. O., (2001). Linkage disequilibrium mapping of the bolting gene in sea beet using AFLP markers. *Genetics Research*, 77(1), 61–66.
- Hao, Z., Li, X., Xie, C., Weng, J., Li, M., Zhang, D., Liang, X., et al., (2011). Identification of functional genetic variations underlying drought tolerance in maize using SNP markers. *Journal of Integrative Plant Biology*, 53(8), 641–652.
- Hoeschele, I., (2004). Mapping quantitative trait loci in outbred pedigrees. *Handbook of Statistical Genetics*.
- Howarth, C. J. (2005). Genetic improvements of tolerance to high temperature. *Abiotic Stresses: Plant Resistance Through Breeding and Molecular Approaches*. Howarth Press Inc., New York, 277–300.
- Huseynova, I. M., & Rustamova, S. M., (2010). Screening for drought stress tolerance in wheat genotypes using molecular markers. *Proc. ANAS (Biol. Sci.)*, 65(5, 6), 132–139.
- Hussain, B., Lucas, S. J., Ozturk, L., & Budak, H., (2017). Mapping QTLs conferring salt tolerance and micronutrient concentrations at seedling stage in wheat. *Scientific Reports*, 7(1), 1–14.
- Iglesias-García, R., Prats, E., Fondevilla, S., Satovic, Z., & Rubiales, D., (2015). Quantitative trait loci associated with drought adaptation in pea (*Pisum sativum* L.). *Plant Molecular Biology Reporter*, 33(6), 1768–1778.
- Jaikishan, I., Rajendrakumar, P., Madhusudhana, R., Elangovan, M., & Patil, J. V., (2015). Development and utility of PCR-based intron polymorphism markers in sorghum [*Sorghum bicolor* (L.) Moench]. *Journal of Crop Science and Biotechnology*, 18(5), 309–318. doi: 10.1007/s12892-015-0015-y.
- Jain, S. M., (2001). Tissue culture-derived variation in crop improvement. *Euphytica*, 118(2), 153–166.
- Jain, S., Kumar, A., Mamidi, S., & McPhee, K., (2014). Genetic diversity and population structure among pea (*Pisum sativum* L.) cultivars as revealed by simple sequence repeat and novel genic markers. *Molecular Biotechnology*, 56(10), 925–938. doi: 10.1007/s12033-014-9772-y.
- Jeffreys, A. J., Wilson, V., & Thein, S. L., (1985). Hypervariable ‘minisatellite’ regions in human DNA. *Nature*, 314(6006), 67–73.
- Jeon, J. S., Chen, D., Yi, G. H., Wang, G., & Ronald, P., (2003). Genetic and physical mapping of Pi5 (t), a locus associated with broad-spectrum resistance to rice blast. *Molecular Genetics and Genomics*, 269(2), 280–289.
- Jiang, G. L., (2013). Molecular markers and marker-assisted breeding in plants. *Plant Breeding from Laboratories to Fields*, 45–83.
- Jones, E., Sullivan, H., Bhattramakki, D., & Smith, J., (2007). A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (*Zea mays* L.). *Theoretical and Applied Genetics*, 115(3), 361–371.

- Kalia, R. K., Rai, M. K., Kalia, S., Singh, R., & Dhawan, A., (2011). Microsatellite markers: An overview of the recent progress in plants. *Euphytica*, 177(3), 309–334.
- Kang, Y. J., Lee, T., Lee, J., Shim, S., Jeong, H., Satyawani, D., Kim, M. Y., & Lee, S. H., (2016). Translational genomics for plant breeding with the genome sequence explosion. *Plant Biotechnology Journal*, 14(4), 1057–1069.
- Karp, A., (1997). *Molecular Tools in Plant Genetic Resources Conservation: A Guide to the Technologies*. Bioversity International.
- Kebede, H., Subudhi, P., Rosenow, D., & Nguyen, H., (2001). Quantitative trait loci influencing drought tolerance in grain sorghum (*Sorghum bicolor* L. Moench). *Theoretical and Applied Genetics*, 103(2, 3), 266–276.
- Kirigwi, F., Van, G. M., Brown-Guedira, G., Gill, B., Paulsen, G. M., & Fritz, A., (2007). Markers associated with a QTL for grain yield in wheat under drought. *Molecular Breeding*, 20(4), 401–413.
- Klein, A., Houtin, H., Rond, C., Marget, P., Jacquin, F., Boucherot, K., Huart, M., et al., (2014). QTL analysis of frost damage in pea suggests different mechanisms involved in frost tolerance. *Theoretical and Applied Genetics*, 127(6), 1319–1330.
- Koebner, R. M., & Varshney, R. K., (2006). Development and application of genomic models for large-crop plant genomes. *Model Plants and Crop Improvement*, 1(9).
- Kraakman, A. T., Niks, R. E., Van, D. B. P. M., Stam, P., & Van, E. F. A., (2004). Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics*, 168(1), 435–446.
- Kraakman, A., Martinez, F., Mussiraliev, B., Van, E. F., & Niks, R., (2006). Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. *Molecular Breeding*, 17(1), 41–58.
- Kumawat, G., Gupta, S., Ratnaparkhe, M. B., Maranna, S., & Satpute, G. K., (2016). QTLomics in soybean: A way forward for translational genomics and breeding. *Frontiers in Plant Science*, 7, 1852.
- Kumpatla, S. P., Buyyarapu, R., Abdurakhmonov, I. Y., & Mammadov, J. A., (2012). Genomics-assisted plant breeding in the 21st century: Technological advances and progress. In: *Plant Breeding*. Intechopen.
- Lai, Z., Livingstone, K., Zou, Y., Church, S., Knapp, S., Andrews, J., & Rieseberg, L., (2005). Identification and mapping of SNPs from ESTs in sunflower. *Theoretical and Applied Genetics*, 111(8), 1532–1544.
- Lecomte, L., Duffe, P., Buret, M., Servin, B., & Causse, M., (2004). Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. *Theoretical and Applied Genetics*, 109(3), 658–668.
- Li, G., & Quiros, C. F., (2001). Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: Its application to mapping and gene tagging in *Brassica*. *Theoretical and Applied Genetics*, 103(2, 3), 455–461.
- Li, G., Xu, X., Tan, C., Carver, B. F., Bai, G., Wang, X., Bonman, J. M., Wu, Y., Hunger, R., & Cowger, C., (2019). Identification of powdery mildew resistance loci in wheat by integrating genome-wide association study (GWAS) and linkage mapping. *The Crop Journal*, 7(3), 294–306. doi: <https://doi.org/10.1016/j.cj.2019.01.005>.
- Li, Z., & Nelson, R. L., (2002). RAPD marker diversity among cultivated and wild soybean accessions from four Chinese provinces. *Crop Science*, 42(5), 1737–1744.

- Liang, X., Chen, X., Hong, Y., Liu, H., Zhou, G., Li, S., & Guo, B., (2009). Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and *Arachis* wild species. *BMC Plant Biology*, 9(1), 35. doi: 10.1186/1471-2229-9-35.
- Lin, K. H., Lo, H. F., Lee, S. P., George, K. C., CHEN, J. T., & Yeh, W. L., (2006). RAPD markers for the identification of yield traits in tomatoes under heat stress via bulked segregant analysis. *Hereditas*, 143(2006), 142–154.
- Linh, L. H., Linh, T. H., Xuan, T. D., Ham, L. H., Ismail, A. M., & Khanh, T. D., (2012). Molecular breeding to improve salt tolerance of rice (*Oryza sativa* L.) in the Red River delta of Vietnam. *International Journal of Plant Genomics*, 2012.
- Liu, B. H., (2017). *Statistical Genomics: Linkage, Mapping, and QTL Analysis*. CRC press.
- Liu, B., Kanazawa, A., Matsumura, H., Takahashi, R., Harada, K., & Abe, J., (2008). Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene. *Genetics*, 180(2), 995–1007.
- Liu, R., Fang, L., Yang, T., Zhang, X., Hu, J., Zhang, H., Han, W., et al., (2017). Marker-trait association analysis of frost tolerance of 672 worldwide pea (*Pisum sativum* L.) collections. *Scientific Reports*, 7(1), 1–10.
- Mago, R., Spielmeyer, W., Lawrence, G., Lagudah, E., Ellis, J., & Pryor, A., (2002). Identification and mapping of molecular markers linked to rust resistance genes located on chromosome 1RS of rye using wheat-rye translocation lines. *Theoretical and Applied Genetics*, 104(8), 1317–1324.
- Mammadov, J., Aggarwal, R., Buyyarapu, R., & Kumpatla, S., (2012). SNP markers and their impact on plant breeding. *International Journal of Plant Genomics*, 2012.
- Massman, J., Cooper, B., Horsley, R., Neate, S., Dill-Macky, R., Chao, S., Dong, Y., et al., (2011). Genome-wide association mapping of fusarium head blight resistance in contemporary barley breeding germplasm. *Molecular Breeding*, 27(4), 439–454. doi: 10.1007/s11032-010-9442-0.
- Miles, C., & Wayne, M., (2008). Quantitative trait locus (QTL) analysis. *Nature Education*, 1(1), 208.
- Mohammadi-Nejad, G., Arzani, A., Reza, A., Singh, R., & Gregorio, G., (2008). Assessment of rice genotypes for salt tolerance using microsatellite markers associated with the saltol QTL. *African Journal of Biotechnology*, 7(6).
- Mondal, S., Badigannavar, A. M., & D'Souza, S. F., (2012). Development of genic molecular markers linked to a rust resistance gene in cultivated groundnut (*Arachis hypogaea* L.). *Euphytica*, 188(2), 163–173. doi: 10.1007/s10681-011-0619-3.
- Mondini, L., Noorani, A., & Pagnotta, M. A., (2009). Assessing plant genetic diversity by molecular tools. *Diversity*, 1(1), 19–35.
- Moose, S. P., & Mumm, R. H., (2008). Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology*, 147(3), 969–977.
- Moreau, L., Charcosset, A., Hospital, F., & Gallais, A., (1998). Marker-assisted selection efficiency in populations of finite size. *Genetics*, 148(3), 1353–1365.
- Munns, R., & Tester, M., (2008). Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol.*, 59, 651–681.
- Najeb, B. M., Al-Doss, A. A., Elshafei, A. A., & Moustafa, K. A., (2011). Identification of new microsatellite marker linked to the grain filling rate as indicator for heat tolerance genes in F2 wheat population. *Australian Journal of Crop Science*, 5(2), 104.

- Nicot, N., Chiquet, V., Gandon, B., Amilhat, L., Legeai, F., Leroy, P., Bernard, M., & Sourdille, P., (2004). Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). *Theoretical and Applied Genetics*, 109(4), 800–805.
- Nirgude, M., Babu, B. K., Shambhavi, Y., Singh, U. M., Upadhyaya, H. D., & Kumar, A., (2014). Development and molecular characterization of genic molecular markers for grain protein and calcium content in finger millet (*Eleusine coracana* (L.) Gaertn.). *Molecular Biology Reports*, 41(3), 1189–1200. doi: 10.1007/s11033-013-2825-7.
- Nordborg, M., Borevitz, J. O., Bergelson, J., Berry, C. C., Chory, J., Hagenblad, J., Kreitman, M., et al., (2002). The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics*, 30(2), 190–193.
- Oropeza, M., & Degarcia, E., (1997). Use of molecular markers for identification of sugar cane varieties (*Saccharum* sp). *Phyton-International Journal of Experimental Botany*, 61(1, 2), 81–85.
- Parh, D. K., Jordan, D. R., Aitken, E. A. B., Mace, E. S., Jun-ai, P., McIntyre, C. L., & Godwin, I. D., (2008). QTL analysis of ergot resistance in sorghum. *Theoretical and Applied Genetics*, 117(3), 369–382. doi: 10.1007/s00122-008-0781-8.
- Park, Y. H., Alabady, M. S., Ulloa, M., Sickler, B., Wilkins, T. A., Yu, J., Stelly, D. M., et al., (2005). Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred line cotton population. *Molecular Genetics and Genomics*, 274(4), 428–441.
- Poortavakoli, S., Sheidai, M., Alishah, O., & Noormohammadi, Z., (2017). Genetic diversity analysis in drought stress tolerant cottons. *The Nucleus*, 60(1), 57–62.
- Popi, J., Rajnpreht, J., Kannenberg, L., & Pauls, K., (2000). Random amplified polymorphic DNA-based evaluation of diversity in the hierarchical, open-ended population enrichment maize breeding system. *Crop Science*, 40(3), 619–625.
- Pupilli, F., Labombarda, P., Scotti, C., & Arcioni, S., (2000). RFLP analysis allows for the identification of alfalfa ecotypes. *Plant Breeding*, 119(3), 271–276.
- Qiu, L., Yang, C., Tian, B., Yang, J. B., & Liu, A., (2010). Exploiting EST databases for the development and characterization of EST-SSR markers in castor bean (*Ricinus communis* L.). *BMC Plant Biology*, 10(1), 278. doi: 10.1186/1471-2229-10-278.
- Rafalski, J. A., (2002). Novel genetic mapping tools in plants: SNPs and LD-based approaches. *Plant Science*, 162(3), 329–333.
- Ragot, M., & Hoisington, D., (1993). Molecular markers for plant breeding: Comparisons of RFLP and RAPD genotyping costs. *Theoretical and Applied Genetics*, 86(8), 975–984.
- Rashed, M., Sabry, S., Atta, A., & Mostafa, A., (2010). Development of RAPD markers associated with drought tolerance in bread wheat (*Triticum aestivum*). *Egyptian Journal of Genetics and Cytology*, 39(1).
- Ribaut, J. M., & Betrán, J., (1999). Single large-scale marker-assisted selection (SLS-MAS). *Molecular Breeding*, 5(6), 531–541.
- Roy, J. K., Smith, K. P., Muehlbauer, G. J., Chao, S., Close, T. J., & Steffenson, B. J., (2010). Association mapping of spot blotch resistance in wild barley. *Molecular Breeding*, 26(2), 243–256. doi: 10.1007/s11032-010-9402-8.
- Sabouri, H., & Sabouri, A., (2008). New evidence of QTLs attributed to salinity tolerance in rice. *African Journal of Biotechnology*, 7(24).

- Saeed, M., Wangzhen, G., & Tianzhen, Z., (2014). Association mapping for salinity tolerance in cotton ('*Gossypium hirsutum*' L.) germplasm from US and diverse regions of China. *Australian Journal of Crop Science*, 8(3).
- Saleh, B., (2016). DNA changes in cotton (*Gossypium hirsutum* L.) under salt stress as revealed by RAPD marker. *Advances in Horticultural Science*, 30(1), 13–22.
- Santos, F. R. C., Pinto, L. R., Carlini-Garcia, L. A., Gazaffi, R., Mancini, M. C., Gonçalves, B. S., Medeiros, C. N. F., et al., (2015). Marker-trait association and epistasis for brown rust resistance in sugarcane. *Euphytica*, 203(3), 533–547. doi:10.1007/s10681-014-1257-3.
- Sardesai, N., Kumar, A., Rajyashri, K., Nair, S., & Mohan, M., (2002). Identification and mapping of an AFLP marker linked to Gm7, a gall midge resistance gene and its conversion to a SCAR marker for its utility in marker aided selection in rice. *Theoretical and Applied Genetics*, 105(5), 691–698.
- Sasanuma, T., Miyashita, N., & Tsunewaki, K., (1996). Wheat phylogeny determined by RFLP analysis of nuclear DNA. 3. Intra-and interspecific variations of five *Aegilops* sitopsis species. *Theoretical and Applied Genetics*, 92(8), 928–934.
- Sbei, H., Sato, K., Shehzad, T., Harrabi, M., & Okuno, K., (2014). Detection of QTLs for salt tolerance in Asian barley (*Hordeum vulgare* L.) by association analysis with SNP markers. *Breeding Science*, 64(4), 378–388.
- Shao, H. B., Liang, Z. S., Shao, M. A., & Sun, Q., (2005). Dynamic changes of anti-oxidative enzymes of 10 wheat genotypes at soil water deficits. *Colloids and Surfaces B: Biointerfaces*, 42(3, 4), 187–195.
- Shasany, A., Darokar, M., Dhawan, S., Gupta, A., Gupta, S., Shukla, A., Patra, N., & Khanuja, S. P., (2005). Use of RAPD and AFLP markers to identify inter-and intraspecific hybrids of *Mentha*. *Journal of Heredity*, 96(5), 542–549.
- Siedler, H., Messmer, M., Schachermayr, G., Winzeler, H., Winzeler, M., & Keller, B., (1994). Genetic diversity in European wheat and spelt breeding material based on RFLP data. *Theoretical and Applied Genetics*, 88(8), 994–1003.
- Sonah, H., Bastien, M., Iquira, E., Tardivel, A., Légaré, G., Boyle, B., Normandeau, É., et al., (2013). An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PloS One*, 8(1), e54603.
- Soto-Cerda, B. J., Saavedra, H. U., Navarro, C. N., & Ortega, P. M., (2011). Characterization of novel genic SSR markers in *Linum usitatissimum* (L.) and their transferability across eleven *Linum* species. *Electronic Journal of Biotechnology*, 14(2), 4. <http://dx.doi.org/10.2225/vol14-issue2-fulltext-6>.
- Spaniolas, S., May, S. T., Bennett, M. J., & Tucker, G. A., (2006). Authentication of coffee by means of PCR-RFLP analysis and lab-on-a-chip capillary electrophoresis. *Journal of Agricultural and Food Chemistry*, 54(20), 7466–7470.
- Suenaga, K., Singh, R., Huerta-Espino, J., & William, H., (2003). Microsatellite markers for genes Lr34/Yr18 and other quantitative trait loci for leaf rust and stripe rust resistance in bread wheat. *Phytopathology*, 93(7), 881–890.
- Tadesse, W., Suleiman, S., Tahir, I., Sanchez-Garcia, M., Jighly, A., Hagra, A., Thabet, S., & Baum, M., (2019). Heat-tolerant QTLs associated with grain yield and its components in spring bread wheat under heat-stressed environments of Sudan and Egypt. *Crop Science*, 59(1), 199–211.
- Tanksley, S. D., Young, N. D., Paterson, A. H., & Bonierbale, M., (1989). RFLP mapping in plant breeding: New tools for an old science. *Bio/Technology*, 7(3), 257–264.

- Upadhyaya, H. D., Wang, Y. H., Sharma, R., & Sharma, S., (2013). Identification of genetic markers linked to anthracnose resistance in sorghum using association analysis. *Theoretical and Applied Genetics*, 126(6), 1649–1657. doi: 10.1007/s00122-013-2081-1.
- Varshney, R. K., Kudapa, H., Pazhamala, L., Chitikineni, A., Thudi, M., Bohra, A., Gaur, P. M., et al., (2015). Translational genomics in agriculture: Some examples in grain legumes. *Critical Reviews in Plant Sciences*, 34(1–3), 169–194.
- Wainaina, C. M., Makihara, D., Nakamura, M., Ikeda, A., Suzuki, T., Mizukami, Y., Nonoyama, T., et al., (2018). Identification and validation of QTLs for cold tolerance at the booting stage and other agronomic traits in a rice cross of a Japanese tolerant variety, Hananomai, and a NERICA parent, WAB56-104. *Plant Production Science*, 21(2), 132–143.
- Wan, H., Chen, L., Guo, J., Li, Q., Wen, J., Yi, B., Ma, C., et al., (2017). Genome-wide association study reveals the genetic architecture underlying salt tolerance-related traits in rapeseed (*Brassica napus* L.). *Frontiers in Plant Science*, 8, 593.
- Wang, Z., Li, J., Luo, Z., Huang, L., Chen, X., Fang, B., Li, Y., Chen, J., & Zhang, X., (2011). Characterization and development of EST-derived SSR markers in cultivated sweetpotato (*Ipomoea batatas*). *BMC Plant Biology*, 11(1), 139. doi: 10.1186/1471-2229-11-139.
- Watanabe, S., Xia, Z., Hideshima, R., Tsubokura, Y., Sato, S., Yamanaka, N., Takahashi, R., et al., (2011). A map-based cloning strategy employing a residual heterozygous line reveals that the GIGANTEA gene is involved in soybean maturity and flowering. *Genetics*, 188(2), 395–407.
- Weiland, J., & Yu, M., (2003). A cleaved amplified polymorphic sequence (CAPS) marker associated with root-knot nematode resistance in sugar beet. *Crop Science*, 43(5), 1814–1818.
- Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., & Tingey, S. V., (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, 18(22), 6531–6535.
- Williams, K., (2003). The molecular genetics of disease resistance in barley. *Australian Journal of Agricultural Research*, 54(12), 1065–1079.
- Wright, S. I., Bi, I. V., Schroeder, S. G., Yamasaki, M., Doebley, J. F., McMullen, M. D., & Gaut, B. S., (2005). The effects of artificial selection on the maize genome. *Science*, 308(5726), 1310–1314.
- Wu, R., Ma, C., & Casella, G., (2007). *Statistical Genetics of Quantitative Traits: Linkage, Maps and QTL*. Springer Science & Business Media.
- Xiao, Y., Pan, Y., Luo, L., Zhang, G., Deng, H., Dai, L., Liu, X., et al., (2011). Quantitative trait loci associated with seed set under high temperature stress at the flowering stage in rice (*Oryza sativa* L.). *Euphytica*, 178(3), 331–338.
- Xu, M., Xu, Z., Liu, B., Kong, F., Tsubokura, Y., Watanabe, S., Xia, Z., et al., (2013). Genetic variation in four maturity genes affects photoperiod insensitivity and PHYA-regulated post-flowering responses of soybean. *BMC Plant Biology*, 13(1), 1–14.
- Xu, Y., (2010). *Molecular Plant Breeding*. CABI.
- Xue, D., Huang, Y., Zhang, X., Wei, K., Westcott, S., Li, C., Chen, M., et al., (2009). Identification of QTLs associated with salinity tolerance at late growth stage in barley. *Euphytica*, 169(2), 187–196.

- Yan, J., Kandianis, C. B., Harjes, C. E., Bai, L., Kim, E. H., Yang, X., Skinner, D. J., et al., (2010). Rare genetic variation at *Zea mays crtRB1* increases β -carotene in maize grain. *Nature Genetics*, 42(4), 322–327.
- Yan, J., Warburton, M., & Crouch, J., (2011). Association mapping for enhancing maize (*Zea mays* L.) genetic improvement. *Crop Science*, 51(2), 433–449.
- Zabeau, M., (1993). *Selective Restriction Fragment Amplification: A General Method for DNA Fingerprinting*. European Patent Application 92402629.7.
- Zhang, G. L., Chen, L. Y., Xiao, G. Y., Xiao, Y. H., Chen, X. B., & Zhang, S. T., (2009). Bulk segregant analysis to detect QTL related to heat tolerance in rice (*Oryza sativa* L.) using SSR markers. *Agricultural Sciences in China*, 8(4), 482–487.
- Zhang, J., Song, Q., Cregan, P. B., Nelson, R. L., Wang, X., Wu, J., & Jiang, G. L., (2015). Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. *BMC Genomics*, 16(1), 1–11.
- Zhou, G. A., Chang, R. Z., & Qiu, L. J., (2010). Overexpression of soybean ubiquitin-conjugating enzyme gene GmUBC2 confers enhanced drought and salt tolerance through modulating abiotic stress-responsive gene expression in *Arabidopsis*. *Plant Molecular Biology*, 72(4), 357–367.
- Zhu, C., Gore, M., Buckler, E. S., & Yu, J., (2008). Status and prospects of association mapping in plants. *The Plant Genome*, 1(1).



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Index

A

- Abiotic stresses, 17, 271, 274
- Absorbent material, 176
- Abzymes, 24
- Accidental leakage, 229
- Acetaldehyde, 227
- Acetohydroxyacid synthase, 227
- Acetoin, 226–228
- Acetylation, 184
- Achromobacter lipolyticum*, 226
- Acidaminococcus* sp., 225
- Acinetobacter*
 - baumannii* S30, 229
 - sp. BS3, 230
 - venetianus strain, 229
- Acrylamide, 178, 204, 208
- Active phosphorus, 33
- Acylation, 184
- Addition
 - primary antibody, 184
 - secondary antibody, 184
- Adenine, 68, 98, 154, 161, 163, 175
- Adenosine deaminase (ADA), 19
- Adenoviruses, 78, 139
- ADP-ribosylation, 184
- Adsorption, 84, 125, 214, 234
- Aerobacter*, 226
- Agarose, 90, 175, 204, 206, 207, 212
 - concentration, 175
 - gel, 115, 135, 174–176, 204, 206–208, 212, 260–263, 266
- Agricultural
 - characteristics, 269
 - operations, 235
 - production, 51
- Agrobacterium*, 18, 79, 86
 - methods, 85
 - tumefaciens*, 79, 86, 87
- Agronomic field, 265
- AIDS causative agent, 138
- Alcohol precipitation, 171
- Alkaline phosphatase (ALP), 12, 59
- Alkylating, 184
 - agents, 114
- Alleles, 128, 135–137, 158–161, 252, 255, 259, 261, 269, 273
 - mutants, 160
- Allergic research, 51
- Allolactose, 102
- Amino acid, 53
 - sequence, 56, 153
- Amniocentesis, 139, 140
- Amniotic fluid, 139
- Ampicillin, 6, 101, 103
- Amplicon, 124, 125–127, 129–132, 134
- Amplification, 29, 34, 66, 86, 119, 121, 123, 124, 128, 129, 131–138, 142, 207, 260, 262, 263, 265, 266
 - DNA, 66, 119, 266
 - fragment length polymorphism (AFLP), 137, 256, 258, 262, 263, 265
 - long templates, 136
 - rate, 86
 - reaction, 124, 135
 - refractory mutation system (ARMS), 128, 142
- Amylases, 11
- Amyloidogenesis, 240
- Anabolic metabolism intermediates, 53
- Analysis, 190, 204, 209, 214, 215
 - detection, 119
 - diversity, 136
 - paternity, 136
- Anaphoresis, 200
- Anionic plasma membrane, 84
- Annealing
 - assisted ligation, 66
 - temperature, 123, 124, 126, 127, 131
- Anoxic condition, 239
- Anthraquinone dye, 232

Anthrax, 141
 Anti-biofilm enzyme gene, 238
 Antibiotic
 formation, 215
 modification, 24
 resistance, 5, 15, 62, 63, 101, 106, 138, 238
 gene, 62, 63, 78, 101, 106
 tobacco plant, 52
 Antibodies, 50, 65
 molecules, 50
 Anti-convective nature, 212
 Antigens, 50, 86, 178
 Anti-hemophilic factors, 50
 Antimicrobial strategy, 238
 Antioxidants, 53
Arabidopsis thaliana, 60, 158, 159
 Arginylation, 184
Aspergillus, 236
 niger, 223, 228, 230
 oryzae, 35, 223
 Association mapping, 252–255, 273
 Asymmetric fashion, 27
 ATP-dependent phosphodiester bond, 33
 Atrazine
 biodegradation, 237
 chlorohydrolase, 237
 contaminated soils, 237
 Autofluorescing components, 132
 Autoimmune diseases, 239
 Autonomous
 replicating
 nature, 4
 sequence (ARS), 108, 116
 self-replicating behavior, 5
 Autoradiography, 19, 173, 184, 259
 Autosomal chromosome, 195
 Auxiliary proteins, 98
 Avian myeloblastosis virus (AMV), 29, 60
 Azo dye, 232

B

Bacillus, 236
 latrosporus RRK1, 232
 stearothermophilus (Bst), 29, 30, 33, 36
 thuringiensis (Bt), 17, 47, 53
 Bacteria, 9, 10, 17, 46, 47, 53, 55, 79, 80, 104, 120, 138, 229, 236, 237
 artificial chromosomes (BAC), 8, 20, 107–109, 116, 174
 vectors, 225
 cell wall, 101
 culture, 47, 223
 defense mechanism, 109
 growth, 101
 host cells, 80
 infections, 173
 sources, 69
 Bacteriophage, 28–30, 34, 35, 47, 58, 59, 64–66, 78, 86, 99, 105, 109, 114, 151, 158, 174, 225
 DNA Vectors, 64
 T4 protein, 59
 B-cell epitopes, 51
 Beta
 galactosidase, 102, 103
 lactam ring (ampicillin), 101
 lactamase, 101
 mercaptoethanol, 209
 Biochemical, 23, 24, 85, 86, 148, 150, 158, 215, 257
 markers, 257
 nature, 215
 rDNA delivery research methods, 86
 reactions, 23, 24
 studies, 85, 148, 158
 Biochemistry, 178, 216
 Biolistic
 gun, 89
 particle delivery, 88
 rDNA delivery, 89
 transformation method, 223
 Biological
 compounds, 16
 pathways, 108
 samples, 122
 scissor, 25, 47
 staining, 210
 Biomedical applications, 69
 Biomolecules, 199–201, 209–212, 214, 216
 Bioprecipitation, 234
 Bioremediation, 220, 230, 231, 234, 236, 240
 Biotin, 125

Biotinylation, 184
Biowarfare agents, 141
Bis-acrylamide, 208
Bisulfite, 224
Blood complications, 234
Blotting, 169, 171, 178, 181
 membrane, 170
 tests, 15
Blue-white screening, 102, 103
Blunt ends, 57
 fragments, 114
Bone marrow transplantation, 19, 140
Boron neutron, 224
Bovine
 pancreas, 34
 serum albumin (BSA), 126, 172, 183, 185
Breeding, 45, 51, 148, 149, 251–253, 258,
 266, 269–273
 performance, 51
 programs, 258
Brewing, 53, 226
Broad-spectrum analyzation, 209
Brucella abortus, 224
Budding, 148
Burkholderia, 236
 seminalis, 224

C

Cadmium, 234
Caenorhabditis elegans, 159, 238
Calcium
 chloride, 80–82, 88
 solution, 13
 phosphate
 co-precipitates, 81
 DNA co-precipitates, 81
 micro-precipitates, 82
 nano-precipitates, 82
 precipitate, 13, 82
Candida glabrata, 227, 228
Capillary
 blotting, 176
 electrophoresis (CE), 211, 213, 214,
 216, 273
Carbamate, 235
Carbohydrate ploy-cationic derivative, 82

Carbonyl cyanide m-chlorophenyl hydra-
 zone (CCCP), 223, 241
Carboxy-fluorescein, 130
Carboxylesterase B1 (CarE B1), 237
Carcinogenic aromatic compounds, 231
Cassette mutagenesis, 224
Cataphoresis, 200
Cationic
 lipid, 78, 81, 83–85, 92
 DNA complex, 84
 rDNA, 83
 liposomes, 83
 polymer, 84, 85
 mediated delivery, 81
Cell
 differentiation, 60
 energy, 221
 materials, 11, 128
 membrane complexity, 223
 mortality ratio, 90
 organelles, 14
Cellulase, 11
 acetate, 211, 213
 electrophoresis, 211, 213
 plate, 213
 acetic acid technique, 213
Cerebrospinal fluids (CSF), 129, 142
Cerrena unicolor BBP6, 230
Chemical
 analysis, 200
 compounds, 53
 mutagenesis, 224
 mutagens, 224
 precipitation, 234
 stability, 179
 transformation, 223
 agents, 223
Chiasmata theory, 150
Chimeric
 DNA, 12
 molecules, 64
 phage, 105
Chlamydia, 139
Chlamydomonas, 236
Chloramphenicol, 101
Chlorella virus, 30
Chloronicotinylnl, 235

- Chorionic villus sampling (CVS), 139, 140
- Chromatids, 150
tetrad, 150
- Chromatography, 173, 175
- Chromium, 234
- Chromogenic detection formation, 173
- Chromosomal, 3, 8, 13, 62–64, 99, 107, 108, 140, 149–151, 153, 154, 157, 159, 160, 162, 189–197, 225, 231, 237, 257, 268, 270, 274
DNA, 61, 62, 104
jumping, 189–192, 194–197
library clones, 192
technique, 189, 194
library, 192, 197
mapping, 195
microarray technique, 195
- Chronological events, 147
- Clinical
diagnosis, 138
research, 169
- Cloning, 4–7, 11, 12, 27, 34–36, 46–49, 59, 61, 63–67, 69, 77, 97–109, 111–115, 119, 124, 141, 171, 172, 189, 190, 192, 194–197, 200, 201, 231, 261, 270
efficiency, 67, 109
genetic markers, 195
vector, 4, 7, 12, 36, 61, 63–66, 97–105, 108, 109, 111–115
- Clostridium*
beijerinckii, 225
thermocellum, 223
- Clustered regularly interspaced short palindromic repeats (CRISPR), 220, 222, 225, 228, 240
Cas technologies, 220
- Cobalt, 234
- Co-dominance, 270
marker, 259, 261
- Co-expressed genes, 159
- Cold season genotypes, 272
- Coliform, 141
- Collagen, 85
- Collimated laser beams, 90
- Collinearity hypothesis, 152
- Commercial
cationic polymers, 85
kits, 139
products, 36
test, 140
- Competent host cell, 115
- Complementary
DNA (cDNA), 11, 19, 34, 35, 59, 60, 66, 68, 105, 120, 134, 135, 264
single-stranded ends, 7
sticky ends, 12, 114
strand, 60, 121
DNA, 60
- Complex prokaryotic replication, 28
- Concanavalin A, 184
- Concatemeric DNA, 66
- Conjugative plasmids, 62, 63
- Contaminants, 129, 230
- Contemporary physical technique, 88
- Conventional
breeding, 252, 253, 268
methods, 139, 140, 234
protocols, 141
techniques, 189
testing, 138, 139
- Coomassie blue, 179, 180
- Copper, 234, 235
- Coronaviruses, 139
- Cosmid, 5–8, 65, 66, 69, 106, 107, 122, 174, 189, 225
vector, 6–8, 65, 66, 69, 106
- Covalent
bonds, 175, 179
closed circular molecules, 57
- CpG dinucleotides, 135
- Creatine kinase, 212
- Cross-pollination, 252
- Crown gall disease, 86
- Cultivar identification, 51
- Culture medium, 12
- Cyclodextrin, 85
- Cynomolgus monkeys, 239
- Cystic fibrosis, 140, 195
disease, 195
- Cytological, 150, 257
markers, 257
- Cytomegaloviruses, 139
- Cytoplasmic
bridge, 13
RNA, 175, 176

Cytosine, 68, 98, 135, 161, 162, 175
Cytosol, 82, 84, 85, 91
Cytotoxicity, 87, 90, 91
 issues, 91

D

De novo

 biosynthetic pathways, 221
 synthesis, 221

DEAE-dextran, 79, 81–83
 mediated approach, 81
 polymer, 82

Decolorization
 ability, 232
 efficiency, 232

Degradation
 efficiency, 229–232
 toluene, 234

Degree of mutability, 160

Deinococcus radiodurans, 234, 235

Deliberation gradient, 202

Denaturation, 68, 120, 131, 133, 171, 172,
 175, 209, 215
 agent, 209, 210

 gradient gel electrophoresis (DGGE), 209

Deoxynucleoside triphosphate substrates, 61

Deoxyribonuclease I (DNase I), 34, 55

Deoxyribonucleic acid (DNA), 1–15, 19,
 20, 25, 27–31, 33–36, 46–48, 51, 52,
 54–69, 77–91, 97–99, 104–115, 117,
 119–127, 129, 131–136, 138, 140, 141,
 151, 153–158, 161–164, 170–174, 176,
 183, 189–197, 200–210, 212, 215, 222,
 224, 225, 251–254, 257, 259–274

 adenylate, 30

 agarose gel electrophoresis, 115

 amplification, 68, 132

 analysis, 56, 141, 189, 197, 206, 215

 delivery method, 83, 86

 deoxyribonucleases, 34

 double helical structure, 151

 electrophoresis, 170

 encoded defense mechanism, 225

 fragments, 12, 14, 57, 63, 65, 68, 69,
 105, 107–109, 111, 113, 114, 140,
 170, 171, 174, 190, 191, 193, 196,
 206, 208

 ligase, 12, 30, 33, 34, 47, 57–59, 63,
 110, 112–115, 191
 microinjection, 90
 moiety, 83
 molecular, 29, 57, 63, 64, 68, 98, 105,
 110, 114, 134, 162, 206, 215
 markers, 51
 polymerase, 28–31, 59, 60, 68, 106,
 124–127, 131, 134
 polymerase I, 28, 29, 60
 protein interaction, 34
 purification, 8, 171
 replication, 5, 30, 35, 60, 98, 113, 196
 sequence, 19, 48, 55, 56, 98, 107, 109,
 153, 154, 173, 192, 195, 254, 257,
 259, 265
 coding, 257
 structural base, 3
 supercoiling, 35
 viral vectors, 86

Deoxyribonucleoside

 chain, 60

 triphosphates (dNTPs), 28, 29, 33, 36,
 60, 120, 123, 125

Deoxyribose sugar molecules, 114

Desulfovibrio desulfuricans, 224

Detection

 genetic disorders, 19

 X syndrome, 140

Diabetes ministration, 239

Diacetyl

 production, 227

 pathway, 227

 reductase (Dr), 170, 227

Diagnosis

 activity, 238

 diseases, 24, 138

 neonatal, 173

Dideoxy

 method, 28

 sequences, 34

Diethylaminoethyl-dextran (DEAE-
dextran), 82, 92

Diffusion, 180, 181

 amplicon, 129

 blotting, 180, 181

 multiple times, 180

Digestion, 29, 34, 35, 54, 57, 112, 113,
115, 128, 171, 178, 190, 191, 211, 259,
261, 263
protein, 34

Dimerization, 102, 103

Dimethyl sulfoxide (DMSO), 83, 92, 125,
126, 209, 223

Di-nucleotide, 260

Dioleoylphosphatidyl ethanolamine
(DOPE), 83, 92

Direct
blue 71 (DB 71), 232
dye, 232
micro-injection, 90

Disperse dye, 232

Dissemination, 47

Divalent cation, 55, 125

D-lactate dehydrogenase (D-LDH), 228

Double
haploid, 255
layered micelle formation, 84
rotational symmetry, 54
stranded
DNA (DsDNA), 29, 33, 34, 54, 58,
60, 61, 63, 67, 134, 170
genomic template, 66
molecule, 110
primer product, 123
RNA, 34

Downward capillary transfer, 176

DreB-B1 gene, 267

Drosophila, 150, 159, 174
experiments, 150
melanogaster, 159
X chromosome, 150

Drought
resistance, 268, 271, 272
sensitivity index, 271
tolerance, 271, 273
associated primers, 267
varieties, 267

Drug-producing *Streptomyces* microbial
species, 89

Dye, 232, 233
degradation, 232

E

Eastern blotting, 183, 184

Ecological synchronization, 240

Edible vaccines, 51

Electric
current, 88
discharge, 89
shocks, 13

Electroblotting, 171, 176, 177, 179–181, 216

Electrochemical treatment, 234

Electrochemistry, 200, 214

Electroelution, 181

Electrokinetic, 201, 216
principle, 199, 201

Electrolytic cell, 201

Electromotive force (EMF), 204

Electronic waste, 234

Electrophoresis, 11, 135, 137, 169–171,
175–177, 181, 197, 199–206, 208–216,
260, 263, 266
chamber, 203, 204
efficiency, 205
gel medium, 202, 205, 206
mobility, 201, 202, 204, 209
movement, 204
process, 203, 204, 206, 208–211
technique, 129, 199, 200, 203, 211

Electroporation, 13, 80, 87, 88, 223, 239

Electrostatic interactions, 84, 85

Embryological intervention, 150

Endocytosis, 81–85

Endonuclease (ER), 9, 11, 12, 23, 25, 34,
47, 54–58, 61, 69, 100, 109–115, 131,
170, 171, 190, 223
enzyme, 11, 34, 171

Endosomal discharge, 81–83, 85

Engineered
bacteria strain, 239
E. coli, 238, 239
microbial strain, 237
strain SYNBI618, 239
strategies, 234
technologies, 219

Enhanced green fluorescent protein
(EGFP), 237

Enterobacter species, 226
aerogenes, 227
cloacae, 227

Enzymatic, 3, 11, 23–25, 27, 28, 34, 36,
47–49, 53–56, 60, 61, 69, 100, 101, 106,
108–110, 112, 123, 127, 152, 163, 180,
190, 210, 222, 223, 236
activities, 69
catalyzed modification, 135
labeled secondary antibody, 183
linked oligonucleotide sorbent assay
(ELOS), 134
renaturation, 181
Epigenetic, 27
inheritance, 157
Epithelial cells of midgut, 17
Epitope detection, 182
Epstein Barr viruses, 139
Escherichia coli, 3, 10, 13, 19, 20, 25, 27,
29, 33–35, 46, 47, 57, 59, 60, 62–66, 88,
141, 158, 191, 223–225, 230, 232, 235,
237–240
DH5 α , 232
Ethidium bromide, 176, 210
migration, 176
Ethyl methane sulfonate, 224
Ethylene-diamine-tetraacetate (EDTA), 10
Euchromatin, 257
Eukaryotes, 5, 12, 30, 60, 67, 78, 80, 84,
88, 104, 109, 122, 154, 155, 157, 158,
162, 163, 196, 215, 220, 221, 240
cells, 12, 61, 78, 80, 88
genes, 163
mRNA precursors, 163
organisms, 30, 221
Exogenous gene, 61
Exonuclease, 28–30, 34, 55, 223
activity, 28–30
Exonucleolytic properties, 61
Exploration, 3, 196, 229
Exponential phase, 53, 131
Expression
studies, 98
vector, 4, 65, 98, 240
Extensive molecular genomic markers, 252
Extrachromosomal
DNAs, 61
double circular DNA elements, 158

F

Fatty acid metabolism, 239
Fenpropathrin, 231, 237
Fermentation, 2, 53, 220, 227, 240
condition, 226
production, 226
Fetal blood cells, 140
Fidelity, 30, 60, 126, 127
enzyme, 30
Field estimation, 266
Filter
membrane type, 177
treatment, 175
Flavr Savr tomatoes, 47, 52
Flocculation, 234
Fluid equilibrium, 50
Fluorescence, 129–133, 142, 184, 210, 214
labeled probes, 129
resonance energy transfer (FRET), 130,
131, 142
signal, 129, 131, 132
Fluorophore, 130, 131
Flux balance analysis, 222, 241
with flux ratios (FBrAtio), 222
Food
Drug Administration (FDA), 10, 52
globalization, 220
industries, 53, 220
ingredient production, 53
pharmaceutical sectors, 227
processing industry, 226
production, 225, 240
security, 52
Forensic, 27, 140
studies, 173
Formaldehyde, 175, 176
Formamide, 125, 175
Formylation, 184
Frost-tolerance genotypes, 270
Fungal
mycelium, 223
resistance, 69

G

Galactose, 102, 103, 206
Gel
chamber, 204

- electrophoresis, 128, 134, 170, 171, 174, 175, 178–180, 183, 184, 189, 200, 202, 203, 210–212, 215, 216, 260, 262, 263
- Gene
 - cloning, 52, 56, 69, 108, 113, 192
 - detection, 255
 - expression, 24, 51, 67, 83–85, 97, 163, 238
 - gun, 14
 - approach, 88
 - machine, 11
 - rearrangement, 173
 - replication, 5
 - specific second primer complementary, 134
 - transcript, 156, 163
- Generally recognized as safe (GRAS), 226
- Genetic
 - analysis, 137
 - circuit, 238
 - code, 9, 152, 153
 - combinations, 48
 - composition, 265
 - counseling, 169
 - differences, 266, 270
 - disorders, 128, 139, 140, 189, 196
 - diversity, 136, 137, 256, 265–268, 273
 - disease, 2, 19, 27, 69, 140, 173, 195
 - engineered (GE), 1–3, 12, 15, 16, 18, 19, 23, 25, 27–30, 33, 36, 45–48, 52, 53, 87, 104, 219–222, 225–227, 229–241
 - crops, 52
 - microbes, 236
 - microorganisms (GEM), 219, 220, 231–234, 236–238, 240, 241
 - evaluation, 269
 - factors, 272
 - fingerprinting, 215
 - information, 151, 159, 237
 - makeup, 15, 20, 49
 - manipulation, 46, 50, 224, 225, 238
 - marker, 159, 189, 194–197
 - material, 1, 2, 4, 8, 9, 19, 46, 49, 51, 52, 67, 87, 88, 97, 148, 151, 158, 161, 220
 - modification (GM), 1, 10, 17, 46, 49–53, 59, 69, 141, 220, 222, 226, 230, 234–237, 240, 241, 266
 - organisms (GMO), 1, 20, 27, 46, 47, 52, 221, 222, 224, 229, 240, 241
 - predeposition, 220
 - recombination, 9, 150, 151, 224, 225
 - screening (hybrids), 265
 - strategies, 222
 - structure, 5, 45, 253
 - studies, 67, 265
 - transformed organisms, 229
 - variation, 97, 255, 258, 266, 268–271, 273, 274
- Genic molecular markers, 259
- Genome, 46, 48, 108, 173, 225, 254
 - abnormality, 215
 - abundance, 262
 - alteration, 221
 - analysis, 189
 - DNA, 11, 67, 106, 120, 122, 171, 224, 259–263
 - libraries, 106
 - libraries, 66–68, 190
 - screening, 67
 - sequence, 10, 224, 251, 262, 264
 - template stability, 266, 274
 - variations, 195
 - wide population, 255
- Genotypic, 255, 262, 263, 271, 274
 - molecular marker, 255
 - utilization, 255
- Geranylgeranylation, 184
- Germplasm, 51, 255, 258, 267
- Giemsa stain, 257
- Glass
 - injection needle, 14
 - microfluidic chip, 182
- Global environmental stresses, 267
- Glucagon-like peptide 1 (GLP-1), 239, 240
- Glucose, 102, 213, 228, 239
- Glucosyl-hydroxymethylated, 27
 - DNA, 110
- Glutamylation, 184
- Glutathione S-transferases (GST), 236
- Glycerol, 125
- Glycoconjugates, 182
 - detection, 183

Glycosylation, 184
Glyoxal, 175, 176, 209
 gels, 176
Green fluorescent protein, 4, 237
Guanine, 68, 98, 161, 175

H

Haemophilus influenzae, 25, 55
Hairpin loop, 126
 structure, 130
Haplotypes, 254
Heat
 resistance, 268
 tolerance, 268, 269
 cultivars, 268
Heavy metal
 contamination, 234
 toxicity, 234
Helicobacter pylori, 223
Hemophilia A, 50
Hemophilia B gene therapy, 50
Hemophilus pleuropneumoniae, 223
Hepatitis B, 10, 51
 virus (HBV), 51, 69
Hepatitis C virus, 139
Herbicide, 235
 tolerance, 52
Herpes
 simplex virus (HSV), 129, 142
 viral vectors, 86
Heterochromatin, 257
Heterogeneous family, 135
Heterolactic acid bacteria, 226
Heterozygous, 160, 259, 261
High
 cost consumption, 91
 efficiency rate, 90
 fructose corn syrup (HFCS), 53
 performance
 capillary electrophoresis (HPCE), 213
 liquid chromatography (HPLC), 213,
 214
 resolution melting, 128
 voltage electrical impulses, 13
Himulin, 18
Homogeneous
 gene sequences, 225
 product, 11
 sample, 216
Homopolymer tail, 66, 114
 mediated ligation PCR, 66
Homozygous, 160, 259, 261
Horizontal
 blotting, 181
 gene transfer (HGT), 55, 222, 241
Host
 cell toxicities, 86
 parasite interactions, 56
 plant defense, 18
 specificity, 86, 222
Human
 alpha-lactalbumin, 16
 environmental protection, 52
 gene therapy, 49
 immunodeficiency virus, 60
 papillomaviruses, 138
 parvoviruses, 139
 protein-rich milk, 16
 Serum Albumin (HSA), 50
 T-lymphocyte viruses, 138
Hybrid oligonucleotide, 66
Hybridization, 19, 68, 121, 124, 128, 130,
 131, 134, 136, 140, 141, 170–173, 259
 allele specific oligonucleotide probe,
 140
 chamber, 172
 florescent probe, 173
 probe, 130, 131
Hydrazine, 224
Hydrocarbon concentration, 229
Hydrogen
 bonds, 133, 162
 peroxide, 266
Hydrolysis, 56, 59, 101, 131
 penicillin, 101
Hydrophobic
 region, 85
 tail, 83
Hydroxyl radicals, 266
Hydroxymethylated, 27, 110
 DNA, 27
Hypertonic solution, 91
Hypoallergenic vaccines, 51
Hypotonic solution, 91

I

Imidazole
 acetol phosphate ester transaminase, 161
 glycerol phosphate ester dehydrase, 161
Immune
 electrophoresis, 211
 system, 49, 138, 159
 cells, 49
Immunoblotting, 178, 180
Immunoglobulin gene, 156
Immunoglobulin, 156, 174
 genes manipulation, 50
In vitro
 cell free experiments, 151
 condition, 33, 238
 environment, 65
 methods, 5
 recombinants, 225
 situations, 266
Incubating cell culture, 171
In-depth exploration, 3
Indigoid dye, 232
Industrialization, 232
Infirmary prenatal diagnosis, 196
Inflammatory
 conditions, 240
 responses, 240
Ingenious modifications, 65
Inhibitors accumulation, 126
Insecticidal Bt genes, 53
Insertion, 1, 65, 97, 103, 154, 224, 227,
 229, 230, 232, 239, 259, 263, 264
 sequences-transposons, 154
Insulin, 18
Intercalation, 176
Interferon, 47
Intermediate-tolerant lines, 270
Intermolecular interaction, 209
International Chicken Polymorphism Map
 Consortium, 137
Interracial
 evolutionary relations, 136
 population, 137
Interrogation, 192
Intestinal biomarker, 239
Intramolecular strand ligation, 34
Intron, 152, 156, 163, 196
 free DNA copies, 60

 sequence, 154
Invertebrates, 86
In-vitro assembling, 106
Ionic
 interaction, 82, 84, 176
 strength, 123, 125
Iso-electric
 focusing (IEF), 180, 210, 211, 214
 histidine, 205
Isoenzyme mixture, 34
Isoprenylation, 184
Isopropyl β -d-1-thiogalactopyranoside
 (IPTG), 102, 116
Isoschizomers, 25
Isotachopheresis (ITP), 211, 214, 216

J

Jumping library, 189, 192–196
Juxtaposition, 130

K

Kanamycin, 101
K-casein, 16
Klebsiella, 236
 pneumoniae, 223
Kohlrausch equation, 200

L

Lactate dehydrogenase (Ldh), 212, 227
Lactic acid
 bacteria (LAB), 226, 227
 tolerant yeast, 228
Lactobacillus, 225–228
 casei, 227, 228
 plantarum, 227, 228
 reuteri, 225, 227, 228
Lactococcus, 226, 228, 239, 240
 lactis, 226, 228, 239, 240
LacZ gene, 102, 103
Lambda, 7, 64, 65
 DNA, 7
 phage, 7, 12, 65, 122
 cos sequence, 7
Laser
 induced fluorescence (LIF), 214
 mediated transfection, 90, 91

microscopic systems, 91
 Leguminous
 crops, 254
 family, 254
 Leishmaniasis, 139
 Lethal methylmercury hydroxide, 209
Leuconostoc, 226, 228
 Leukemias, 134
 Ligase, 58
 adenylate, 30
 enzyme, 30, 33
 polymerases, 36
 Ligation, 12, 34, 54, 59, 66, 107, 112–114,
 191, 192, 263
 oligonucleotide, 34
 sticky end fragments, 114
 Linear
 DNA molecule, 202
 fragment method, 225
 red system, 225
 regression, 132
 sequential arrangement, 154
 Lipases, 11
 Lipids, 11, 13, 83, 182, 183
 Lipofection, 13
 Lipoplex, 84
 Lipoylation, 184
Listeria monocytogenes, 141
 Lithium borate LB buffer, 205
 Log-linear
 curve, 132
 phase, 132, 133
 Long
 non-coding RNA (lncRNA), 157, 164
 terminal repeats (LTRs), 67
 Lymphocytes, 19
 Lysosomal hydrolases, 83
 Lytic development, 106

M

Macromolecules, 13, 53, 200, 204, 205, 212
 Macroscopic techniques, 182
 Magnesium, 27
 Mammalian
 artificial chromosome (MACs), 8, 20
 cell, 8, 60, 80, 86, 88, 89, 155, 171
 cultures, 81
 genome, 197
 host cells, 82
 Manganese, 231, 234
 peroxidase (MnP), 230, 231
 Mannitol production, 228
 Mapping mammalian genomes, 107
 Marker, 4, 51, 62, 91, 136, 137, 160,
 190–192, 194, 195, 200, 215, 252, 253,
 255–261, 264–274
 assisted breeding (MAB), 251–254, 273,
 274
 formation, 215
 Mass spectroscopy, 180
 Maternal
 blood, 140
 sample, 140
 pattern (inheritance), 136
 Matured phage particle, 66
Meloidogyne incognita, 18
 Membrane
 bound glycoprotein, 59
 filtration, 234
 permeability, 13
 Mercury, 234, 235
 Metabolic
 engineering, 221, 222, 226, 234
 tools, 222
 flux, 221, 222
 management, 221
 strategies, 220, 222
 Metallic oxides, 199, 216
 Metallome, 210
 Metallothionein (MT), 235
 Metaplasia, 135
 Methyl
 group, 35, 54, 135
 parathion, 237
 hydrolase gene (Mpd), 237
 Methylase, 23, 54
 enzyme, 35
 Methylated, 27, 55, 109, 135, 184
 DNA, 135
 Methyltransferase coding gene, 28
 Microbes
 bioremediation process, 230
 cell, 50, 223
 factories (MCF), 221, 241
 genes, 50

- communities, 236
 - consortia, 230
 - degradation, 235
 - hosts, 89
 - insecticides, 53
 - populations, 230
 - products, 53
 - recombination, 13
 - remediation, 232
 - Microbiology, 128, 220
 - Microbiota, 237, 238
 - Microchip, 182
 - electrophoresis (MCE), 214
 - Microfluidic analysis platform, 182
 - Micro-needles, 90
 - Micro-projectile
 - approach, 88
 - bombardment, 14
 - gun, 14
 - Microsatellites, 136, 137, 270
 - Mineralization, 236
 - Minisatellites, 136
 - Mitochondria, 14, 89, 91, 158, 184
 - DNA, 120, 136, 141
 - Mixed function oxidases (MFO), 236
 - Molecular
 - analysis, 201, 214
 - beacons, 130, 131
 - biology, 2, 23, 28, 48, 58–60, 77, 82, 83, 87, 92, 141, 149, 164, 169, 170, 190, 200, 236, 252
 - breeding, 252, 260, 263, 265
 - composition, 201
 - genetics, 274
 - linkers, 28
 - markers, 205, 251–253, 255, 257, 258, 264–271, 273, 274
 - mass (protein), 178, 204
 - proteins, 180
 - scissors, 109
 - techniques, 137
 - transformations, 266
 - understanding (inheritance), 48
 - weight, 61, 178–180, 191, 203–205, 212, 213
 - Monogenic trait, 252
 - Monovalent cations, 125
 - Moraxella* sp., 236
 - Mucopolysaccharides, 122
 - Multi-allelic
 - markers, 273
 - nature, 269
 - Multicellular eukaryotes, 157
 - Multi-genetic forms, 269
 - Multiple cloning site (MCS), 7, 65, 99, 100, 102, 103, 116
 - Mutagenesis, 59, 141, 222, 224
 - Mutation
 - breeding technology, 253
 - mapping, 35
 - rates, 109
 - Mycobacteria* sp., 223
 - tuberculosis*, 138, 139
 - Mycoplasma*
 - genitalium*, 138, 158
 - mycoides*, 27, 28
- N**
- National Center for Biotechnology Information, 126
 - Native
 - gel, 210
 - microflora, 231, 240
 - Natural
 - bio-structure of molecules, 210
 - cationic polymers, 85
 - ecosystem, 231
 - flora, 231
 - mechanism (infection), 87
 - microflora, 230, 232
 - polymers, 85
 - sugars, 87
 - Negative cell membrane, 84
 - Neoclassical, 148, 151–153, 155, 156
 - period, 151, 152
 - Neoplasia, 135
 - Neoschizomers, 25
 - Nervous system, 234
 - Nested PCR, 135
 - Nitroalkylation, 184
 - Nitrocellulose, 170–172, 179, 180, 183, 184, 259
 - filter, 172

- membrane, 170–172, 179, 180, 183, 184
 - powder, 170
 - Nitrogen
 - bases, 161, 202
 - fixation efficiency, 254
 - Nitrophosphomolybdate-methyl green, 184
 - Nitrous acid, 224
 - Non-conjugative plasmids, 63
 - Non-essential
 - elements, 105
 - genes, 105
 - Non-GM sugar beets, 52
 - Non-Hodgkins lymphoma, 134
 - Non-homologous chromosomes, 157
 - Non-integrating plasmids, 5
 - Non-palindromic forms, 27
 - Non-specific
 - amplification, 123
 - binding, 123, 125, 177, 179
 - product, 123–126
 - Non-target organisms, 235
 - Non-thermostable DNA ligases, 58
 - Non-transformed cells, 15
 - Non-uniformly charged test species, 205
 - Non-viral
 - chemical means, 80
 - DNA
 - delivery systems, 82
 - delivery vectors, 86
 - transfer, 83
 - gene transfer methods, 78
 - vectors, 50, 85
 - Northern blotting, 174–178
 - Novel
 - drug formulation, 220
 - solicitation, 216
 - Novozyme, 223
 - Nuclease, 34, 54, 173
 - enzyme category, 25
 - protection, 177
 - Nucleic acid, 3, 13, 34, 54, 59, 69, 85, 90, 91, 134, 169, 171, 177, 199–202, 204, 205, 209–212, 215, 216, 259
 - proteins, 169
 - Nucleophilic attack, 33
 - Nucleoside monophosphates, 59
 - Nucleosome assembly, 36
 - Nucleotide, 28, 29, 35, 53, 57, 59, 68, 98, 111, 122, 123, 129, 154, 161, 162, 260, 262
 - arrangement, 110
 - sequence, 54, 55, 64, 97, 109, 110, 127, 152, 171, 257
 - Nucleotidyl transferase, 61
 - Nutritional
 - composition, 226
 - sweetener, 53
 - Nylon, 179
 - membranes, 171, 177
- O**
- Okazaki fragments, 30
 - Oligodeoxynucleotide, 61
 - Oligomer formation, 123
 - Oligonucleotide, 34, 135
 - directed mutagenesis, 224
 - Oncogenic genes, 87
 - Open reading frames (ORF), 153, 163, 164
 - Operon genes, 164
 - OptForce, 222
 - Optimization, 89, 121
 - PCR methodology, 128
 - OptKnock, 222
 - Opto-poration, 90, 91
 - cell membrane, 91
 - Oral Phe dietary supplementation, 239
 - Organelle membranes, 184
 - Organic
 - acids, 53
 - matter, 230
 - Organochlorines, 235
 - Organophosphates, 235
 - Organophosphorus (OP), 236
 - Origin of replication (Ori), 4, 6, 62, 98, 99, 102, 104, 106, 109, 116
 - sequence, 98, 109
 - Oryza sativa*, 253
 - Osmotic difference, 91
 - Oxidative destruction, 266
- P**
- P1-derived artificial chromosomes (PACs), 8, 20, 174
 - Palindrome, 25, 54, 56, 110
 - sequence, 25, 27, 56, 110

- Pancreas, 159
 - ribonuclease, 34
- Paper electrophoresis, 211
- Paraffin, 128, 172
- Parallel
 - mapping strategies, 197
 - sequencing, 66
- Parent genotypes, 252, 262
- Particle
 - bombardment, 79, 88
 - method, 88, 89
 - gun
 - device, 89
 - method, 88
- Paternity testing, 27, 265
- Pathogenic microbes, 238
- Pattern
 - inheritance, 149, 159
 - mutability, 160
- Pectinase*, 11
- Pediococcus*, 226
- Peptide
 - cationic polymers, 85
 - mass fingerprinting, 211
- Permeability, 13, 83, 89
- Permease, 163
- Pesticide, 235, 236
 - degradation, 236
 - bacteria, 236
 - destruction, 236
- Petroleum
 - contaminated soil, 231
 - hydrocarbon, 229
 - products, 229
- Phagemids, 65, 122
- Phagocytosis, 13
- Phanerochaete chrysosporium*, 231, 236
- Pharmaceutical, 15, 17, 97, 231
 - companies, 49
- Pharmacogenetics, 128
- Pharmacogenomics, 49, 69
- Phenanthrene, 230, 231
- Phenanthrene
 - biodegradation, 231
 - degradation, 230
 - capacity, 230
- Phenotypic, 61, 160, 253, 255–257, 271, 273
 - properties, 255
 - traits, 61
- Phenylalanine (Phe), 239
 - metabolizing enzyme encoding genes, 239
- Phosphatase, 54
 - activity, 59
- Phosphate, 13, 28, 30, 33, 54, 55, 58, 59, 79, 81, 82, 110, 113, 161, 202, 214
 - containing substrates, 55
 - linkage, 202
 - monoesters, 59
- Phosphodiester, 25, 33–35, 54, 57, 58, 113, 114, 161
 - bond, 25, 33–35, 54, 58, 113, 161
- Phospholipid bilayers, 88
- Phosphopantetheinylation, 184
- Phosphoramidate, 30, 33
- Phosphorylation, 59, 184
- Photopolymerization, 182
- Photo-transfection, 90, 91
- Phylogenetic, 136, 141, 173, 252, 255, 256, 270, 273
- Physarum*
 - polycephala*, 35
 - polycephalum*, 35
- Physical
 - gene delivery approaches, 90
 - methodologies, 78, 85
 - mutagenesis, 224
- Pichia pastoris*, 230
- Pisum sativum*, 264, 271, 272
- Plant
 - breeding, 48, 253, 257, 267, 274
 - cell, 12, 14, 80, 87, 89–91, 267
 - bioreactors, 53
 - protoplasts, 13, 14
 - tissue culturing, 266
 - viral vectors, 86
- Plaque screening, 65
- Plasma
 - HSA (pHSA), 50
 - membrane, 13, 83, 84, 184
 - oncotic pressure, 50

- Plasmid, 5–8, 12, 19, 34, 46, 47, 61–63, 65, 66, 69, 82, 83, 87, 99, 104–107, 109, 112, 122, 158, 174, 189, 225, 229, 231, 239
borne characteristics, 62
containing transformants, 62
size, 104
vectors, 5, 7, 69, 225
- Plasminogen activator, 17
- Plasmodia, 49
- Pnitrophenol (PNP), 236
- Poly
adenylation, 154
protein genes, 155
- Poly-A polymerase, 35
enzyme, 35
- Polyacrylamide, 170, 176, 178, 180, 185, 203, 204, 206, 208, 212, 214, 216, 259
gel (PAGE), 170, 176, 178–180, 185, 203–205, 208–210, 212, 216, 259
electrophoresis, 216, 259
- Polybrene, 85
- Polycistronic mRNA, 164
- Polycyclic aromatic hydrocarbon, 230
- Polydeoxynucleotide, 61
- Polydeoxyribonucleotides, 60
- Poly-dT purification, 35
- Polyethylene glycol, 33
- Polyethyleneimine, 85
- Polygenic, 252
- Polylinker sequence, 100
- Polylysine, 85
- Polymerase, 19, 23, 28–30, 35, 47, 54, 59, 60, 68, 69, 120, 123–126, 142, 163
chain reaction (PCR), 10, 11, 19, 29, 30, 34, 60, 66–69, 115, 119–129, 131–142, 172, 173, 215, 224, 225, 259–265, 269
reaction mixture, 68
site-directed mutagenesis, 224
enzyme, 28–30, 68
- Polymerase I, 28, 29
- Polymerase II, 28
- Polymerase III, 28
- Polymerization, 28–30
rate (T4 polymerase), 29
- Polymorphic, 135, 255, 258–260, 262, 264, 266, 268, 269
gene family, 135
primer, 266
- Polynuclease enzyme, 28
- Polynucleotide, 161
chain, 162
phosphodiester connection, 55
kinase (Pnk), 59
- Polyornithine, 85
- Polyploidy cells, 160
- Polyprotein gene, 155
- Polysaccharide, 11, 213
cationic polymers, 85
- Polyvinylidene fluoride (PVDF), 179, 181, 183, 185
- Population genotyping, 261
- Positive lipoplex, 84
- Post
amplification process, 129
transcriptional, 153
base modification, 153
translated
modifications (PTMs), 80, 109, 182, 184, 185
proteins expression, 184
- Poverty reduction, 52
- Pre-existing DNA-genome, 190
- Pre-hybridization solution, 172
- Primary metabolites, 53
- Primer
designing, 263
dimers, 123, 126
- Production
antibodies, 156
blue colonies, 103
clotting factors, 47
- Products discrimination, 130
- Progeny, 61, 148, 149, 260, 266, 270, 271
- Prokaryotes, 5, 162, 163, 190, 197
cells, 80, 88
systems, 109
- Proliferator-activated receptors α (PPAR α), 239
- Proof-reading activities, 29
- Propionibacterium*, 226

Protein, 8, 9, 11, 17, 18, 46–49, 56, 59, 99,
109, 151, 153, 155, 159, 163, 164, 169,
171, 178–184, 208, 211, 212, 215, 216,
221, 238, 257
analyzation, 208
coding transcripts, 157
encoding genes, 159
engineering, 3
fingerprinting, 201
glycosylation, 183
immobility, 178
molecules, 24, 207, 208, 215
moonlighting, 153
production, 46
protein interactions, 159
purification, 210
sequences, 169
synthesis
 efficiency, 159
 machinery, 151
Proteolytic
 cleavage, 155
 preparation, 29
Proteomics, 210
Protoxins, 17
Pseudomonas, 230, 231, 236
 , 223, 224, 231, 235, 238
 infection, 238
 GP3A, 230
 putida, 229–231, 234, 235
 KT2440, 231, 235
 strain BNF1, 229
Pulse
 field
 gel electrophoresis (PFE), 208
 gradient electrophoresis, 203
 generator apparatus, 88
Purification process, 106
Pyrethroid, 235
 hydrolase gene (pytH), 231
 insecticides, 237
Pyrimidine, 161
 purine variations, 264
Pyrococcus furiosus, 125
Pyrosequencing, 128

Q

Qualitative traits, 252
Quantitative
 analysis, 130
 PCR, 129
 properties, 269
 trait locus, 254, 274
Quinacrine hydrochloride, 257
Quorum sensing (QS), 238

R

Radioactive, 19, 259
 isotopes, 124
 labeling, 141
 markers, 125
 probes, 29
Real time (RT), 67, 121, 129–135, 138
 instrumentation, 132
Recognition sequences, 55, 56
Recombinant, 10, 14, 15, 18, 23, 27, 45,
46, 50, 58, 65, 69, 77–80, 82, 85, 86, 92,
98, 104, 105, 114, 229–232, 234, 235,
255, 265, 268, 271, 272
Acinetobacter
 baumannii S30 pJES, 229
 BS3-C230, 230
cells, 15, 104
DNA (rDNA), 3–5, 10, 12, 14, 15, 18,
19, 27, 45–49, 51–53, 58, 61, 63,
69, 77–81, 83–92, 97, 98, 105, 110,
113–115, 236
 molecule, 4, 61, 63, 91
 technology, 3–5, 10, 18, 46–49, 51,
52, 69, 83, 84, 86–88, 97, 105, 110,
113, 114
HSA (rHSA), 50
hybrid lines population, 272
molecules, 65
plasmid DNA moiety, 82
Pseudomonas putida KT2440-rhLABRI,
231
Streptomyces coelicolor M145, 230
Redox potential, 230
Remediation processes, 232
Replication, 8, 28, 46, 57, 59, 63–67, 77,
98, 99, 107, 120, 121, 123, 224
 machinery, 98

- Resistant wheat genotypes, 266
 - Restriction
 - digestion, 27, 259–261
 - endonucleases, 6, 12, 25, 36, 47, 57, 100, 170
 - enzymes, 25, 27, 35, 47, 54–57, 64, 65, 69, 100, 109–111, 124, 137, 191, 261
 - nomenclature, 54
 - fragment, 54, 56, 63
 - length polymorphism (RFLP), 54, 128, 137, 256–259, 264, 272
 - modification, 55
 - site, 25, 27, 54, 57, 63, 100, 109, 110, 112, 115
 - Retroviral, 60, 67, 86, 138
 - vectors, 86
 - Reverse
 - exonuclease activity, 28
 - orientation, 112
 - osmosis, 234
 - transcriptase, 29, 60, 67, 134
 - enzyme, 60
 - PCR (RT-PCR), 60, 67, 129–135, 138
 - transcription, 11, 67, 120
 - Reversible phosphorylation, 59
 - Rhizoctonia solani*, 224
 - Rhodothermus marinus*, 33
 - Ribonuclease, 11, 35
 - Phy M, 35
 - Ribonucleic acid (RNA), 2, 3, 9, 11, 18, 24, 25, 33–36, 48, 49, 59, 60, 67, 99, 120, 133, 134, 138, 141, 152–154, 157, 158, 161–164, 170, 172, 174–177, 183, 200, 203–206, 209, 212, 215, 234
 - denaturation, 175
 - fractionation, 174, 175
 - interference, 18
 - polymerase, 152, 162, 163
 - sequencing, 35
 - Ribozymes, 24
 - Ring spot virus, 52
 - Rotaviruses, 139
- S**
- Saccharomyces cerevisiae*, 60, 174, 225, 228
 - S-adenosyl-methionine, 35, 56
 - Saline
 - affected areas, 266
 - sodium citrate (SSC), 172, 185
 - stress, 265–267
 - tolerance, 270, 271
 - genes, 266
 - varieties, 265
 - Salmonella*, 141, 160, 223, 234, 240
 - enterica*, 234
 - typhimurium*, 223, 240
 - Salt
 - sensitive varieties, 266
 - stress tolerance, 252
 - tolerant
 - sequences, 266
 - transmission, 271
 - Saltol introgression, 271
 - Sauerkraut*, 226
 - Secondary
 - metabolites, 254
 - waste products, 234
 - Selectable
 - marker gene, 66, 78, 87, 101, 108, 109
 - representation (epitopes), 51
 - Sequence
 - complementary, 68
 - marker, 263
 - nucleotides, 3
 - Serratia marcescens*, 224
 - Severe combined immunodeficiency (SCID), 49
 - Sexual reproduction, 148
 - Shewanella halifaxensis* HAW-EB4, 240
 - Shockwave mediated transformation, 223
 - Sickle cell anemia, 140
 - Simian virus 40 (SV40), 4, 46, 57, 66, 83, 156
 - monkey DNA virus genome, 4
 - Single
 - base substitution, 126
 - microfluidic platform, 182
 - nucleotide polymorphism (SNP), 127, 136, 142, 252, 254, 256, 258, 263, 264, 272, 273
 - sequence repeat (SSR), 252, 256–258, 260, 264, 268–270, 272, 273
 - strand
 - conformation polymorphism (SSCP), 137, 210, 272

- DNA molecule, 106, 114
 - duplex DNA, 33
 - RNA, 34
 - transcription unit, 155
 - Singlet oxygen, 266
 - Site directed mutagenesis, 224
 - Small regulatory RNAs (sRNA), 221
 - Smooth polysaccharide capsule, 161
 - S-nitrosylation, 184
 - Sodium
 - alginate immobilization surfaces, 90
 - bisulfite, 135
 - dodecyl sulfate (SDS), 10, 170, 172, 178–181, 185, 205, 209, 210, 215, 216
 - polyacrylamide gel electrophoresis, 205
 - phosphate solution, 82
 - Solubility, 230, 231
 - Somaclonal variation, 265
 - Somatic recombination, 156
 - Sorbitol, 227
 - production, 227
 - Spectrophotometer, 176
 - Sphingobium yanoikuyae*, 224
 - Standard
 - biological protocols, 106
 - genetic parameters, 137
 - Streptococcus*, 225, 226, 228
 - pneumonia*, 225
 - Streptomyces*, 223, 236
 - fradiae*, 223
 - Stress
 - resistance, 237, 252, 274
 - tolerance, 267–269, 271, 274
 - Stringency, 123, 125
 - regulation, 99
 - Structural
 - documentation (DNA), 2
 - elucidation, 2
 - mutation, 267
 - Sub-cellular DNA transfer, 91
 - Sub-optimization, 129
 - Substrate nucleotide selectivity, 60
 - Succinylation, 184
 - Sulfuric acid, 213
 - Superoxide, 266
 - Synthetic
 - dyes, 231
 - linkers, 66
 - peptides, 51
- T**
- Taq polymerase, 30, 60, 120, 122, 125, 126, 131
 - Temperature
 - gradient gel electrophoresis (TGGE), 209
 - range, 120, 127
 - Template, 68
 - Tetrachloro-6-carboxyfluorescein, 130
 - Tetramethyl-6-carboxyrhodamine, 130
 - Tetrathionate, 239, 240
 - Thalassemia, 140
 - Therapeutic, 18, 24
 - agent, 237, 238
 - agrochemical field, 220
 - engineered bacteria, 238
 - sector, 220
 - Thermal
 - conventional force, 212
 - cycle, 68, 128, 129, 140
 - denaturation, 34
 - Thermococcus litoralis*, 125
 - Thermodynamics, 222
 - Thermolabile group, 124
 - Thermophilic, 120
 - Thermophilus aquaticus*, 30, 60
 - Thermostable
 - DNA
 - ligases, 59
 - polymerases, 60
 - ligase, 33
 - Thermus*
 - aquaticus*, 60, 120, 125
 - scotoductus*, 33
 - thermophiles*, 33, 59
 - Thioredoxin, 29
 - Thymine, 68, 98, 161, 162
 - Ti-plasmid, 87
 - Tissue specific expression, 170
 - Topoisomerase, 35, 36
 - Tradition
 - cotton, 52

- food products, 240
- methods, 138, 231, 234
- treatment methods, 232
- Trait
 - dependent marker link, 269
 - linked polymorphism, 255
- Trans-cinnamate*, 239
- Transcription, 152, 153, 222
 - inhibition, 135
 - inhibitors, 175
 - initiation site, 162, 163
 - proteins, 162
- Transduction, 5, 49, 66, 77–79, 85, 90, 222
- Transfection, 12, 77–79, 83, 86, 89
- Transformation, 6, 12, 13, 46, 62, 77–82, 87–89, 91, 102, 103, 105, 109, 141, 151, 161, 222, 223, 253
 - efficiency, 78, 88, 223
 - inefficiency, 6
- Transfusion medicines, 140
- Transgenesis, 61
- Transgenic
 - animals, 15, 17
 - host, 18
 - mice, 16
 - plants, 51, 52
 - technology, 51
- Transglutamination, 184
- Transmission
 - energy, 130
 - hereditary, 148
 - inheritance pattern, 148
- Transposition mutagenesis, 224
- Transversion mutation, 224
- Treatment
 - charge syndrome, 196
 - hemophilia, 47
 - human diseases, 50
 - infectious disease, 238
- Treponema pallidum*, 138
- Trichoderma* spp., 236
- Triphenyl methane dye, 232
- Tris-acetate-EDTA (TAE), 205, 206
- Tris-borate-EDTA (TBE), 205
- Tris-hydroxymethyl aminomethane, 211
- Triticum aestivum*, 253, 264
- Tuberculosis, 138
- Tungsten, 14, 89
 - particles, 14
- U**
- Ubiquitination, 184
- Unbranched
 - cationic polymers, 85
 - polymers, 85
- Unculturable microflora, 108
- Uniprimer, 131
- Universal
 - mechanism (operation), 147
 - quencher, 130
- Un-methylated cytosine, 135
- Urbanization, 231, 232
- UV measurements, 214
- UV radiation, 114, 171, 224
- V**
- Vaccine, 50
 - production, 50
 - technology, 51
- Vacuum blotting, 179
- Vibrio*
 - anguillarum*, 224
 - cholerae*, 238
 - fischeri*, 224
- Viral
 - pathogens, 138
 - vectors, 5, 6
 - methods, 85
- Viruses, 5, 6, 12, 19, 30, 64, 85, 86, 139, 158, 161
- W**
- Warm alkali treatment, 173
- Watson Cricks model, 170
- Western blotting, 178
- Wheat genetic variation, 258
- White rot mushrooms, 236
- Whole-genome sequencing, 254
- Winter-tolerance, 270
 - associated markers, 270
 - genotypes (peas), 270

X

Xenopus laevis, 154

X-ray induced changes, 150

Y

Yatalase, 223

Yeast

artificial chromosome (YAC), 8, 9, 20,
108, 109, 116

telomere, 108

Yellowstone National Park bacteria, 30

Yogurt, 225–227

Z

Zea mays, 191, 253, 264, 274

Zinc, 33, 234

Zone electrophoresis, 211, 212

Zygote, 14

Zymography, 180